

University of Groningen

Individual differences in very young children's English acquisition in China

Sun, He; Steinkrauss, Rasmus; Tendeiro, Jorge; de Bot, Kees

Published in:
Bilingualism: Language and Cognition

DOI:
[10.1017/S1366728915000243](https://doi.org/10.1017/S1366728915000243)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2016

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Sun, H., Steinkrauss, R., Tendeiro, J., & de Bot, K. (2016). Individual differences in very young children's English acquisition in China: Internal and external factors. *Bilingualism: Language and Cognition*, 19(3), 550-566. <https://doi.org/10.1017/S1366728915000243>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Individual differences in very young children's English acquisition in China: Internal and external factors*

HE SUN

University of Groningen (RUG), the Netherlands

RASMUS STEINKRAUSS

University of Groningen (RUG), the Netherlands

JORGE TENDEIRO

University of Groningen (RUG), the Netherlands

KEES DE BOT

University of Groningen (RUG), the Netherlands

(Received: July 28, 2014; final revision received: April 27, 2015; accepted: April 27, 2015; first published online 3 June 2015)

This study assesses the impact of internal and external factors on very young EFL learners in an instructional setting. 71 child English learners in China (onset age: 2;0 - 5;6) were involved: their receptive vocabulary, productive vocabulary and receptive grammar were taken as outcome variables, and age of onset, short-term memory, nonverbal intelligence, English input quantity and quality, English use, and maternal English level were taken as predictive variables. Multiple regression analyses, verified by Bayes factor comparisons, revealed that the total amount of school input and home English media environment were significant predictors for all of three aspects of English proficiency, with each aspect having different additional significant predictors. Both internal factors (e.g., age of onset) and external factors (e.g., English input quantity) played an important role, but in contrast to similar studies (e.g., Paradis, 2011) focusing on a L2 naturalistic setting, external factors explained more variance of English proficiency measures.

Keywords: internal and external factors, EFL, ESL, early foreign language teaching and learning, Chinese learners

1. Introduction

Over the past few years, research on child second language (L2) acquisition has shown a growing interest in various factors that might influence L2 children's linguistic development (Unsworth, Hulk & Marinis, 2011). Both internal factors, such as age and short-term phonological memory, and external factors, such as quantity and quality of input, have been claimed to affect the rate, route, and end state of L2 learning (e.g., Paradis, 2011). Understanding the impact of different factors on L2 development could not only bring insight into existing theories (e.g., Universal Grammar vs. Usage-Based Theory), but could also enable researchers to formulate pedagogical advice for educators and parents.

The target children in early L2 learning studies are usually from immigrant families who live in a L2 speaking country and need to learn the L2 to communicate in daily life. They are learning the L2 in a naturalistic setting. However, it is still an open question to what extent the findings on these children's L2 acquisition apply to child foreign language (FL) acquisition in an instructional setting. Child FL learners refer to those learners who are

learning another language as part of their curriculum, but do not have to use it very often since they do not live in an L2 dominant context. Compared with child L2 learners in naturalistic settings, the typical child FL learners in instructional settings usually have some, or all, of the following features as summarized by Muñoz (2008): 1) they have limited L2 exposure in general, and the instructional time in class is approximately 50 minutes per week; 2) the quantity and quality of the L2 in class is heavily influenced by the teachers' L2 proficiency and the amount of L2 use in class; 3) the L2 is not adopted by the peers during children's communication; and 4) the use of the L2 outside the classroom is rare. Therefore, child L2 learners in either a naturalistic or an instructional setting might show a significant variance in learning outcome due to different L2 environments.

Out of various early FL programs all over the world, the early English programs in East Asia have developed particularly fast (Butler, 2013). China, for example, has the world's largest number of children learning English as a foreign language (EFL) in instructional settings (Ministry of Education of the People's Republic of China, 2001). Moreover, in recent years, the increasing number of bilingual kindergartens and private English institutes enables millions of Chinese children as young as 2 to 3 years old to start learning English (Sun, de Bot & Steinkrauss, in press). Despite the huge numbers, very little research has been done on these young EFL

* The authors are indebted to Sharon Unsworth and Linda Visser for their comments on an earlier version of this manuscript. The authors would also like to thank Yong Sun and Jinrong Wang for collecting data, Richard Morey for discussing Bayes factor analyses, and Linda Postma and Dylan Armstrong for proofreading the paper.

Address for correspondence:

He Sun, Graduate School for the Humanities, Department of Applied Linguistics, University of Groningen, Rode Weeshuisstraat 12, 9712 ET Groningen, PO Box 716, 9700 AS Groningen, the Netherlands

h.sun@rug.nl

learners. The current study takes previous findings, especially Paradis' findings (2011) on factors impacting the acquisition of children's English as a second language (ESL) in a naturalistic setting, into account and examines the impact of internal and external factors on child EFL learners in China.

2. Internal and external factors in very early L2 learning

The onset age of child L2 learners in instructional settings varies to a great extent. Nikolov and Mihaljević Djigunović (2011) divide these young L2 learners into two groups according to the time of the L2 exposure: children whose ages range from 3 to 6 years old (at preschool) are considered very young learners, and those who are between 7 to 12 years old (at primary school) are considered young learners. This study focuses on very young learners of English in China, including 2-year-olds. The latter is based on the situation in China that many children start their English education at this age. The following sections will introduce findings on the effect of internal and external factors on the L2 acquisition of very young learners of English. Both naturalistic settings and instructional settings are examined.

2.1 Internal factors

Most studies that target very young L2 learners have mainly focused on time-related factors (e.g., age of onset (AoO)) and language aptitude. If the participants come from a heterogeneous population, the transfer of morphosyntactic features from the L1 to the L2 will also be taken into consideration (e.g., Paradis, 2011). Each of these factors could influence children's L2 acquisition rate and ultimate attainment (Unsworth, 2005; Paradis, 2011; Hyltenstam & Abrahamsson, 2003; Unsworth, Persson, Prins & de Bot, 2014).

In terms of AoO, older L2 learners have an initial rate advantage over younger learners in naturalistic settings; however, the younger learners surpass the older learners on ultimate attainment (Muñoz, 2008). For instance, in Jia and Fuse's study (2007), the younger starters mastered the whole morphological system after five years, despite the fact that the older learners demonstrated a steeper growth curve of morpheme learning at the initial stages. However, in an instructional setting, only the initial rate advantage of older learners over younger learners has been proven (García Mayo & García Lecumberri, 2003; Muñoz, 2006), while the ultimate attainment advantage of younger starters over older learners has not (Harley, 1998; Muñoz, 2014). Muñoz (2014) examined the effect of AoO and input on ultimate attainment through the oral performance of 160 non-native speakers of English in Spanish universities. The starting age of the participants

ranged from 3 to 15.5 years old and all participants had learned English for at least 10 years by the time of testing. The results demonstrated that, instead of AoO, input factors (e.g., cumulative exposure and contact with high-quality input) are good predictors of the learners' performance. The author argues that in FL instructional settings, where both input quantity and quality are less favorable than those in naturalistic settings, environmental factors might function as a mediator limiting the effects of AoO on L2 ultimate attainment.

Language aptitude is another important factor that potentially influences child L2 acquisition in both naturalistic and instructional settings. Aptitude refers to the specific capability for language learning which learners are assumed to have (Carroll & Sapon, 2002). It is considered to be relatively stable and composed of several aspects, such as phonemic coding ability, language analytic ability, and memory (Skehan, 1986). The few studies that discuss language aptitude in very young L2 children indicate that short-term memory and analytic reasoning ability are two important components that may predict L2 outcomes in this population (e.g., Alexiou, 2009; Genesee & Hamayan, 1980; Paradis, 2011; Unsworth et al., 2014); good memory and analytical ability seem to facilitate both vocabulary and grammar acquisition.

2.2 External factors

External factors, or environmental factors, have also been found to be significantly related to child SL learners' acquisition, particularly in instructional settings (Muñoz, 2014). These factors include input quantity, input quality, home English environment, parents' L2 proficiency, social economic status (SES), and use of English at home (Unsworth, 2013a).

In terms of input quantity, child L2 learners usually demonstrate a different overall length of exposure (LoE) and current exposure time at preschool and at home (Paradis, 2011). Input quantity has been found to influence not only vocabulary acquisition (Ojima, Matsuba-Kurita, Nakamura, Hoshino & Hagiwara, 2011; Oller & Eilers, 2002; Vermeer, 2001; Unsworth et al., 2014), but also grammatical knowledge (Chondrogianni & Marinis, 2011; Paradis, Nicoladis, Crago & Genesee, 2011). For instance, Ojima et al. (2011) conducted a neuroimaging study in order to address the effects of AoO and input quantity on child FL learners' semantic processing of spoken English. For this, 350 Japanese primary school children (average age at the baseline test = 6–9 years) were included. Results from behavioral and ERP data demonstrated that more hours of exposure to English led to higher English scores, irrespective of whether age of onset was controlled for. The study highlighted the fundamental role of input quantity in children's FL learning. It is worth noting that, traditionally, LoE has been used to estimate the

total amount of input. However, this measurement could be problematic (Stevens, 2006). LoE is the difference between AoO and age at the time of testing (AToT). Once AToT and AoO are held constant (as is often the case in these studies), LoE is constant as well. The correlation of these three factors can make it difficult to accurately identify the cause of the observed effects (Stevens, 2006; Muñoz, 2014). In order to address this issue, Unsworth (2013a) proposed the concept of “cumulative length of exposure” for bilingual children’s studies. She suggested measuring children’s L2 input more accurately by using their daily schedules on a weekday and a weekend (Gutierrez-Clellen & Kreiter, 2003), thereby taking into account differences in the actual amount of input each child receives per year. Consequently, this breaks the fixed relationship between AoO, AToT, and LoE. The current study adopted this idea and measured both the LoE and cumulative exposure in the school setting.

Input quality has been operationalized in a variety of ways. Child ESL learners may differ from each other in terms of input resources (e.g., different types and quantity of media providing English input), parental English ability (usually measured through maternal English proficiency), native input (e.g., number of native English speaking friends), and input intensity (e.g., times of watching English cartoons per week). Rich resources of L2 input (e.g., Jia & Aaronsson, 2003), higher maternal L2 level (e.g., Chondrogianni & Marinis, 2011), more social time with native speakers, and frequent L2 exposure (e.g., Muñoz, 2011, 2014) are all found to facilitate acquisition rates and general outcomes of vocabulary and grammar. Children’s language input at school and at home might be qualitatively different. Therefore, input from different settings should be separately looked into (Cummins, 1984). The current study explored the input factor at school and at home respectively.

Learner’s L2 use, especially the L2 use at home, has also been shown to be a significant predictor for their language acquisition in naturalistic settings. Studies by Paradis (2011), Bohman, Bedore, Pena, Mendez-Perez and Gillam (2010) found that children’s English output is strongly correlated with vocabulary and morphology outcomes. According to Bohman et al. (2010), children’s output is even more important than their input, especially with respect to morphosyntactic acquisition, as practice promotes accuracy and automaticity in the production of grammatical constructions.

Family socio-economic status (SES), usually measured by maternal educational levels in child ESL studies, has also been found to be significantly related to bilingual children’s and very young ESL learners’ English outcome, especially with respect to vocabulary acquisition (e.g., Golberg, Paradis & Crago, 2008; Paradis, 2009; Scheele, Leseman & Mayo, 2010; Hoff, 2006). For instance, Golberg et al. (2008) found that children whose mothers

had post-secondary education had a larger vocabulary size than their peers, whose mothers had secondary-only education. It is worth noting that when SES is measured based on maternal education only, children in the current study seem to come from different classes. However, if SES is measured by Butler’s approach (2013), which not only considers parental education, but also takes their occupations and income into account, most children in the current study seem to come from middle class families.

2.3 Research on very young L2 learners from a multi-factor perspective

While there are studies about the effect of internal and external factors on very young child SL acquisition (see above), few have adopted a comprehensive view to include both aspects and investigate their roles on different language domains, such as vocabulary and grammar, or on the same population (however, see Paradis, 2011, and Chondrogianni & Marinis, 2011). Paradis’ study in 2011 is one of the few studies in which this was done. More importantly, it took a large number of internal and external factors into consideration simultaneously: AoO, AToT, language aptitude (short-term phonological memory and nonverbal intelligence), L1, LoE, Language use at home, English richness, mothers’ English proficiency, mothers’ educational level, and the number of siblings. By investigating 169 children (4;10 to 7;0) from new immigrant families, Paradis found that 1) language aptitude, AToT, L1 typology, LoE, and the richness of an English environment significantly predicted children’s English outcomes; 2) internal factors explained more variance than external ones; and 3) the same factors were found to explain the variance of both the vocabulary and grammatical domains.

To sum up, both internal factors and external factors play a role in child L2 acquisition in naturalistic or instructional settings. Internal factors, such as AoO, short-term memory and nonverbal intelligence, and external factors, such as total amount of input and parents’ L2 proficiency, could impact the outcomes significantly. However, attention should be paid to AoO regarding its different impact on proficiency in different language learning environments. In the long run, the facilitating effect of early AoO, which has been found in naturalistic settings, was not confirmed in studies on instructional settings (Muñoz, 2008). Compared with an early AoO, environmental factors might play a more significant role on the final outcome in instructional settings.

Regarding vocabulary size, older onset age, high language aptitude (short-term memory in particular), sufficient and qualified input, and frequent language use could facilitate its development. Regarding grammatical knowledge, early AoO, good language aptitude, similar grammatical constructions between L1 and L2, good

input quantity and quality, higher maternal L2 proficiency, and L2 use could promote its acquisition. According to Paradis' research, internal factors are more crucial for the speed of L2 acquisition in a naturalistic setting. The question that remains in the research reported here is to what extent this also holds for very young FL acquisition in instructional settings.

2.4 Research Questions and hypotheses

Based on the findings mentioned above, the following two questions have been formulated with respect to internal and external factors in very young learners' vocabulary and grammar acquisition in an EFL instructional setting where English is not the community language and English input and use are generally limited.

1. What is the relative contribution of individual internal and external factors to Chinese children's English proficiency?
2. What is the relative contribution of internal and external factors taken as two different sets to predicting variation in children's English proficiency?

For both questions, proficiency in English is operationalized as English receptive vocabulary size, English productive vocabulary size, and English receptive grammar.

For question 1, based on the findings in naturalistic settings, it is expected that both internal and external factors play a role in facilitating English proficiency in very young Chinese EFL learners in instructional settings. AoO, language aptitude, input quantity and quality, and English use could significantly influence both vocabulary and grammar learning.

In relation to question 2, it is expected that external factors, as a whole, might explain at least as much variance in L2 learning outcomes as internal factors, which is in contrast with findings in naturalistic settings. In Paradis' study, on average, the immigrant children were regularly exposed to authentic English for 20 months. The English input came from preschool or school programs. Outside of the classroom, there were many opportunities for children to have access to English, such as talking with their siblings and English speaking peers, playing computer games, watching television, reading books, and participating in organized language related activities. Moreover, the children had the opportunity to practice English both in and out of the classroom. In such a favorable language environment, children's internal factors might have a stronger impact on the L2 outcome. However, in an EFL setting like China, English input is quite limited in general and the actual exposure and usage depends on the efforts made by each family. In the current study, the children's mean length of exposure to

English was equivalent to Paradis' study (2011), namely 20 months. Nevertheless, children in the current study have had only 2 hours of English instruction per week at their English school. Moreover, English exposure at home is scarce compared with that of the children in Paradis' study (2011). English was neither used in the community, nor adopted by their peers. There were few organized activities in English. In terms of English media use, families varied greatly from each other. Therefore, differences in input quantity, quality, and usage from child to child might have a strong influence on their learning outcomes in the current instructional setting.

3. Methods

3.1 Participants

The participants of this study were 71 Chinese children from Happy English, a private English language institute in Chongqing, China. Happy English is one of the largest private English initiation language institutes in the south-east part of China, focusing on children whose ages range from two to twelve years old. The textbooks for very young language learners in this school are from the Yippee series (Red, Green, and Blue), published by MM publications. They are designed for very young children who learn English as a foreign language, targeting their listening and speaking ability at the beginning and gradually focusing on their reading and writing ability as they get older. A Total Physical Response approach (TPR) (Asher, 1996) is used in teaching at Happy English. In TPR, physical movement is used to support verbal input. TPR intends to motivate students to participate more effectively in language activities and reduce their anxiety and stress. In general, children attend Happy English twice a week, taking English classes as extracurricular activities: once for the main class, instructed by a foreign English teacher and a Chinese teacher together, and another time for the activity class, given by a Chinese teacher only. In the main class, children are taught new content, such as words, phrases, and songs, and in the activity class, they review this content and do additional practice. The children spend about two hours at Happy English per week in total. The foreign teachers are from English speaking countries, such as the U.S. and Canada. Moreover, all of them received a TESOL certificate before employment and have been trained by the experienced teachers of Happy English prior to teaching. The Chinese teachers are usually English major graduates with a bachelor degree at least.

The selection criteria for the participants were: 1) they had started to learn English when 2–5 years old and 2) they had no history of language impairment. In the current study, 30 children received all of their formal English input at Happy English. According to the questionnaire, the other 41 participants previously had, or were having,

English classes at their kindergartens while taking classes at Happy English.

3.2 Materials and procedures

Language aptitude tests: short-term memory and analytical reasoning ability

The internal factors that are related to language aptitude were measured through individual, computerized administration of various tests in a quiet room. This session, which includes tests on phonological short-term memory, nonverbal short-term memory, and nonverbal intelligence, lasted for 35 to 40 minutes. During the tests, children were also asked whether they wanted to have a break every 25 minutes. Details of the tests are described below.

Digit span and non-word repetition, the sub-tests of the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen & Rashotte, 1999), were administered to assess the children's short-term phonological memory. The two tests are composed of a list of digits or nonwords in English, which increases in length. They were played to each participant from a computer, and the child was subsequently asked to repeat them immediately after. The scores of the two tasks were summed up, yielding a composite score for short-term phonological memory (VerMem). It should be noted that, to a certain extent, children's phonological memory could be lower when they are measured in their L2 (Thorn & Gathercole, 1999). However, given the feasibility of using children's L2 to test short-term memory (e.g., Unsworth et al., 2014; Paradis, 2011) and the reasonable scores obtained in the current study, this should not influence the data analysis.

The Hand Movements task by the Kaufman Assessment Battery for Children (K-ABC, Kaufman & Kaufman, 1983) was used to measure children's short-term nonverbal memory. It targets children's ability to use visual immediate memory that does not rely on verbal language. The task is composed of a series of hand gestures (fist, palm, and hand), increasing in difficulty. After watching gestures shown on a computer, a child is asked to remember the components and their sequence.

Raven's Colored Matrices (Raven, Raven & Court, 1998, sets A, B and C) were used to tap into children's analytical reasoning, which is assumed to be an important component of language aptitude (Paradis, 2011). For literate L2 learners, this ability is usually measured through a written unfamiliar language (Sawyer & Ranta, 2002). However, for preliterate learners such as those in the current study, nonverbal intelligence (NonInt) was measured as an indicator instead (Genesee & Hamayan, 1980, Paradis, 2011). A colored version of Raven's Standard Progressive Matrices was used and children needed to choose the missing part of a presented pattern from 6 options. In total, there were 36 items that increased in complexity.

In the current study, a small portion of the children were under 5 years old, which is the youngest age suggested by the aptitude tests such as CTOPP. Therefore, the raw score of the assessments was used instead of the standardized score. In the data analysis, chronological age effect has been partialled out from the aptitude performance with regression analysis.

Parental questionnaire, interview, and school online records

A parental questionnaire, an interview, and school online records were used to gather data for several of the internal and all of the external factors per child (see Table 1).

The parental questionnaire (see Appendix I) was designed based on a language background questionnaire of Early Language and Intercultural Acquisition Studies (ELIAS; Kersten, Rohde, Schelletter & Steinlen, 2010) and Utrecht Bilingual Language Exposure Calculator (UBiLEC; Unsworth, 2013a), both of which have been used effectively in several large scale second (foreign) language studies on very young learners in Europe (e.g., Project ELIAS¹ and Project NoRus²). Compared with existing questionnaires, the current one listed more options of English media used at home, namely: English TV programs made in China, English TV programs made in English-speaking countries, English movies, English audio materials, and English materials on electronic devices. These additions are based on fieldwork and interviews with the teachers and parents before launching the project. In the current study, parents had to fill in the questionnaire before the tests started. It was subsequently used for studying children's age of onset, current chronological age, maternal education, the mothers' self-reported English fluency, English input from media at home, and the children's English usage in general (Table 1).

The children's age of onset in relation to English exposure (AoO) was chosen as the moment from which children received systematic and sustained input. Hearing some spoken English from the media or being taught one or two words incidentally was not taken into account. Most children started to learn English either at Happy English or at their bilingual kindergarten. Maternal education (MotEdu) is measured in terms of highest degree; most mothers had obtained a bachelor's degree. Mothers were also asked to rate their oral English proficiency (MotEng) using a five-point scale. The higher they scored, the better their oral English was considered to be. The mean score resulted to be 2.5, indicating that they have limited English communication skills. Out of concern for a potential

¹ Early Language and Intercultural Acquisition Studies <http://www.elias.bilikita.org/>

² Bilingual Norwegian-Russian Children in North Norway: Language Acquisition and Language Use <https://castl.uit.no/index.php/acquisition/research-projects/norus>

Table 1. Raw scores of predictors and outcome variables.

Category	Predictors	Mean	SD	Range	Data source
Internal factors	AoO	45.49	9.79	24–66	Parental report
	AToT	66.00	11.14	42–98	Parental report
	VerMem	15.55	3.33	7–23	CTOPP-digit and nonword tasks
	NonMem	11.75	2.96	5–18	K-ABC-hand gesture task
	NonInte	21.76	6.12	6–33	Raven's
External factors	LoE	19.46	7.90	8–42	School record; Parental report
	SchInput	124.62	63.30	43.33–316.5	Parental report
	HomInput	1.45	1.52	0–9.5	School record; Parental report
	HomMed	3.75	3.68	0–18	School record; Parental report
	MotEng	2.54	.91	1–5	Parental report
	MotEdu	2.76	.89	1–5	Parental report
	EngUse	2.53	1.09	1–5	Parental report
Outcome variables	EngPro	12.14	4.29	2–23	EOWPV-2
	EngRec	22.70	9.42	9–55	PPVT-4
	EngGra	10.94	6.21	1–27	TROG-2

Note. NB: AoO = age of onset in months; AToT = age at the time of testing in months; VerMem = short-term phonological memory score based on digit span and non-word repetition; NonMem = short-term nonverbal memory score based on hand movement repetition; NonInte = non-verbal IQ score as a measure of analytic reasoning; LoE = Length of exposure to English in months; SchInput = total amount of English input at Happy English and bilingual kindergartens in hours; HomInput = weekly English input quantity at home in hours; HomMed = weekly English input quality at home (the sources and frequency of using English media); MotEng = mothers' self-rated proficiency in English on a 1–5 point scale; MotEdu = mothers' highest educational level; EngUse = number of places of using English in total; EngPro = English productive vocabulary size; EngRec = English receptive vocabulary size; EngGra = English receptive grammar skill

negative influence on their children's pronunciation, most mothers are cautious to speak English to their children and usually only do so when helping them to review words learnt in class or when asked by their children to name something in English.

On average, children's home English input (HomInput), including time spent on English media, English books, and games, was 1.4 hours per week (SD = 1.5), indicating there was very limited exposure to English for these children outside of school. English books and games were barely used at home, since parents worry that their own spoken English might negatively influence their children's pronunciation and grammar. The majority of the parents only used English media (e.g., English movies or songs) to maintain contact with English for their children. Many parents selected the material online or at a bookshop. The books that were chosen most frequently were cartoons, like *Dora the Explorer* and *Hello Teddy*. Besides cartoons, applications of accessible English games on both iPad and smart-pen readers were also used to provide English exposure to their children. Most children were attracted by these materials, at least at the beginning, as long as the content was intriguing and the language was easy to comprehend. Their actual English media use varied from family to family: some children could have as much as 9.5 hours of English input per week, while others could have no home English input at all.

Not only did the quantity of home exposure differ per family, but the quality of home input differed as well. There was considerable variation with respect to the number of different types of media and the frequency of use. In the current study, the richness of home media English (HomMed in table 1) was used to capture this quality aspect of home input. Home media English is similar to the factor "English environment richness" that is used in many child ESL studies (e.g., Paradis, 2011), but it is more focused on the number and frequency of home media English input, since most Chinese children have no native English-speaking friends in daily life and, apart from the textbook, barely use any English books themselves. Points were assigned according to whether one specific English media format was used each day and these were totaled to indicate the general richness of media use. The average of home media English richness was 3.75 (1 type of English media and 4.2 times of use per week). Like home input quantity, children vary significantly from each other, ranging from 0 to 18.

Apart from home input, the total input from school settings (SchInput), taking both Happy English and kindergarten into account, has been calculated. As Unsworth (2013a) argued, LoE is a general calculation of input and, therefore, a more accurate number should be generated to indicate the actual amount of language input children have over time. The current study only

checked accumulated input in the school setting, but left out the home setting, since English exposure at home in an EFL environment is probably neither stable over months nor similar to school input in terms of quality. Current home English input has been estimated separately using the parental report. The exact time participants spent learning English at school was estimated through their English class attendance and an interview. The first author checked all the participants' attendance through online attendance system of Happy English for a precise estimate of their class time. If a child was reported to have English at kindergarten, a short interview (Appendix II) was administered to the parents, in which the exact number of hours the child had had English was asked for. By doing so, we were able to get a comparatively precise estimate of the number of input hours for each child in an instructional setting.

Finally, children's English use (EngUse) was measured using the number of settings in which they could systematically use English (e.g., at Happy English and during vacation). Since the amount of time using English varied per child, the number of different settings (0–6), rather than the exact time, was used to estimate the extent of English use. On average, children had 2.5 places to use English: when they were at Happy English, at home, and at kindergarten. The other optional settings were places where children took a vacation, had an English activity, or practiced English, such as the English corner.

English assessments: receptive/productive vocabulary and receptive grammar

Children's English proficiency was operationalized as English receptive vocabulary size, English productive vocabulary size, and English receptive grammar, and it was measured in several tests. These tests were administered individually through a computer in a quiet room. In order to avoid fatigue, all of the English tests were conducted on a different day than the aptitude tests. The examination of English productive vocabulary, English receptive vocabulary, and English receptive grammar lasted for 25 to 30 minutes per child.

The Peabody Picture Vocabulary Test (PPVT-4, Dunn & Dunn, 2007) was used to measure children's receptive vocabulary skill (EngRec). Together with a spoken word, an array of 4 images was presented to children simultaneously. Afterwards, the children were asked to point out the image that best matched the spoken word. For instance, a child saw an array of 4 images: crayons, a cat, a brush, and a sock, and was asked to indicate which one matched the word "cat".

To examine children's English productive vocabulary skill (EngPro), the Expressive One-Word Picture Vocabulary Test-4 (EOWPVT-4, Brownell, 2010) was used. During this test, children were presented with a

picture and they were subsequently asked to name the depicted object, action or attribute in English.

The Test for Reception of Grammar Version 2 (TROG, Bishop, 2003) was used to measure children's receptive grammar skill (EngGra). As in the PPVT, children saw an array of 4 images and heard a spoken sentence at the same time. They were then asked to point out the image that matched the sentence produced by the computer. For instance, a child was shown an array of 4 images: a seated girl, a running cat, a running girl, and a seated cat, and was asked to point out which image goes with the sentence "the girl is sitting".

All three tests have been operationalized with standard procedures. Two examples were demonstrated prior to the tests to make sure that the children understood the requirements.

4. Results

In order to study the influence of internal and external factors on children's early English proficiency we used regression models. English receptive vocabulary, English productive vocabulary, and English receptive grammar were the dependent variables.

First of all, the age effect was partialled out from the nonverbal intelligence, short-term phonological memory, and short-term nonverbal memory test scores. Since a small number of participants were under the minimum age for the tests, the standard scores of the tests could not be used. Instead, the raw scores were put in three simple regression models with AToT as the independent variable, and the standardized residuals were saved as the new values (for details of the decorrelation approach, also see Blom, Paradis & Sorenson Duncan, 2010 and Blom & Paradis, 2014). The standardized residuals were highly correlated with the scores ($r(\text{VerMem} \& \text{VerMemResid}) = .96$; $r(\text{NonMem} \& \text{NonMemResid}) = .85$; $r(\text{NonInte} \& \text{NonInteResid}) = .80$), indicating that the residuals measured the same aspects of cognitive functions as the raw scores. Secondly, the two aspects of short-term memory, phonological short-term memory and nonverbal short-term memory, were averaged to form a composite score of short-term memory (ShoMem).

Next, in order to avoid multicollinearity problems, given our moderate sample size, it was decided to exclude predictors whenever they displayed a moderate to high correlation with other variables ($r = .5-.8$). Non-parametric Spearman's correlations were conducted to check the relationships. Since AToT was moderately correlated with both AoO ($r = .71$) and LoE ($r = .52$), it was not taken into account in the final regression analysis. Home English media and home English input were also highly correlated ($r = .80$) and home English media was kept because parents mentioned that the number of media sources and their weekly use were much more stable than

Table 2. Non-parametric correlations of the 7 predictors in the model analysis.

Factors	1	2	3	4	5	6	7
1 HomMed	–						
2 SchInput	–.02	–					
3 EngUse	.31**	0	–				
4 MotEng	.22	–.09	.40**	–			
5 AoO	–.03	–.14	–.15	–.13	–		
6 NonInte	.05	.01	.18	.09	.01	–	
7 ShoMem	.29*	.04	.13	.12	.05	.29*	–

Note. NB: AoO = age of onset in months; ShoMem = short-term memory score; NonInte = non-verbal IQ score; SchInput = total amount of English input; HomMed = weekly English input quality at home (the sources and frequency of using English media); MotEng = mothers' self-rated proficiency in English;

the time spent on each item. School input and LoE were highly correlated ($r = .79$), and the former was used because it could more accurately reflect children's English input in the instructional setting. Mothers' oral English level, but not mothers' educational level ($r = .62$), were selected for the same reason, since the former could more precisely predict home English quality.

After these reductions, 7 predictors were used in the subsequent analysis: AoO, short-term memory ("ShoMem"), analytical reasoning ("NonInt"), cumulative input in the instructional setting ("SchInput"), current home English media richness ("HomMed"), mothers' oral English proficiency ("MotEng"), and children's English usage ("EngUse").

The non-parametric correlations of these predictors are in Table 2.

To answer RQ1, backward regression analyses were performed to determine which factors best predict children's three aspects of English skills respectively. This approach was chosen to account for possible multicollinearity effects, since some of the predictors were still moderately correlated with each other. It is therefore important to ensure that large portions of explained variance of the dependent variable are not shared between predictors, for the sake of interpretability of the effects and reliability of estimated coefficients.

To answer RQ2, the final predictors identified for each model were divided into two groups (internal factors vs. external factors) and were put in a second hierarchical multiple regression procedure (forced data entry method) in order to reveal how internal factors and external factors as groups contributed to the three aspects of English skills.

The steps mentioned above have been used frequently in other similar studies (e.g., Paradis, 2011; Chondrogianni & Marinis, 2011). The current study was further extended as follows. First, the final models were examined using a cross-validation (CV) analysis to demonstrate their predictive utility on other datasets (Refaeilzadeh, Tang & Liu, 2009). R squared (R^2) values

of multiple regression analyses typically overestimate the population model fit because their computation is too sample-dependent. In contrast, CV allows a more precise description of the predictive value of an estimated model for new (i.e., different) data derived from the same population (Yin & Fan, 2001). It estimates regression models using one part of the data (the training set) and then applies these models to the remaining part of the data (the validation set). In the current study, the leave-one-out procedure was used and 71 regression models were computed based on all subjects except one (for more details, see Refaeilzadeh et al., 2009). Based on these models, predictions could be made for the excluded subject. Since the training set and validation set are disjointed, this provides the chance of testing the model on data external to the ones used to estimate the parameters. Compared to the usual R^2 , the predicted R^2 provided by CV is based on values predicted from models that were estimated independently. Consequently, its value tends to be smaller than the usual R^2 .

Second, Bayes factors (BFs) (Kass & Raftery, 1995) were used to examine which model selected by multiple regressions was the best. Traditionally, in multiple regressions, the best model is selected according to criteria based on statistically significant results, but this actually does not allow the direct comparison between models based on evidence gathered in the data. Problems concerning statistical significance are under discussion (see the "dance" of the p-values argument in Cumming, 2012). BFs avoid possible problems associated with statistical significance with direct model comparisons. This allows not only for assessing which model is better, but also allows for getting information about the extent to which it is better. The BF is the ratio of posterior odds to prior odds of a hypothesis (say H1) over a competing hypothesis (say H2). BFs larger than 1 indicate that the observed data support H1 over H2 (and by how much), and BFs smaller than 1 indicate the reverse. Each model gets a BF in the sense that it is compared to the

Table 3. Backward regression model results for productive vocabulary skill.

	B	SE	β	T	Sig	95% CIs	Part
(Constant)	1.814	2.946		.616	.540	-4.069, 7.697	
HomMed	.357	.124	.306	2.886	.005	.110, .605	.288
SchInput	.015	.007	.216	2.080	.041	.001, .029	.208
MotEng	1.162	.489	.246	2.377	.020	.186, 2.138	.237
AoO	.093	.046	.211	2.026	.047	.001, .184	.202
ShoMem	.655	.284	.241	2.305	.024	.088, 1.223	.230

Note. $R^2 = .352$, $F(5,65) = 7.058$, $p < .001$

Table 4. Backward regression model results for receptive vocabulary skill.

	B	SE	β	t	Sig	95% CIs	Part
(Constant)	9.729	3.056		3.184	.002	3.629, 15.829	
HomMed	.776	.275	.303	2.820	.006	.227, 1.326	.293
SchInput	.032	.016	.213	2.042	.045	.001, .063	.212
EngUse	2.411	.927	.280	2.602	.011	.562, 4.261	.271

Note. $R^2 = .275$, $F(3,67) = 8.480$, $P = .001$

only-intercept (i.e., without predictors) model (Kass & Raftery, 1995, section 3). The advantage of BFs over classical (frequentist) approaches is that BFs allow using evidence as shown by data to quantify the relative adequacy of competing models.

All these steps have been used on the three dependent variables separately, taking into account the same predictor variables each time.

4.1 Regression analyses for productive vocabulary skill

In order to address the first research question on the relative contribution of individual internal and external factors to early language proficiency, all 7 predictors were entered into the backward regression to find the best fitting model. The results are presented in Table 3. The model explained a significant portion of the variance in productive vocabulary skill scores ($R^2 = .352$, $F(5,65) = 7.058$, $p < .001$). The selected predictors were short-term memory, age of onset, mothers' oral English proficiency, total amount of school input, and home English media. Home English media had the largest standardized beta coefficient and semipartial correlation (Beta = .306, semipartial $r = .288$). It can uniquely explain about 8% of the total variance of productive vocabulary skill scores.

Of the five predictors selected by the backward regression procedure, three were external and two were internal. A further step was conducted to answer the second research question that relates to the relative contribution of internal and external factors when taken as two different sets for predicting variation in children's English proficiency. A hierarchical regression analysis

was conducted in which the external factors (SchInput, HomMedia, MotEng) were entered as a first block and the internal factors (AoO and ShoMem) as a second block. Results demonstrated that the external factors alone explained about 26% of the variance in productive vocabulary scores ($R^2 = .258$, $F(3,67) = 7.757$, $p < .001$), and just 9% of additional variance was explained by entering the internal factors (R^2 change = .094, $F(2,65) = 4.719$, $p < .012$). Due to correlations of the internal factors and external factors, this regression was performed again in reversed order: internal factors were entered first, followed by external factors. The result was still that the external factors explained more variance than the internal factors (internal $R^2 = .143$, $F(2, 68) = 5.680$, $p < .005$; external R^2 change = .209, $F(3,65) = 6.979$, $p < .001$).

4.2 Regression analyses for receptive vocabulary skill

As with the productive vocabulary skills, backward regression was used on receptive vocabulary skills (Table 4). The model accounted for 27.5% of the variance in receptive vocabulary skill ($F(3,67) = 8.480$, $p < .001$). All three predictors in the model (total amount of school input, home English media, and English usage) that had significant coefficients ($p < .05$) were external factors. As for the productive vocabulary skill, home English media again had the strongest standardized beta coefficient and partial correlation (Beta = .303, semipartial $r = .293$). As no internal factors were found to be significant predictors for receptive vocabulary skill, no forced entry regression was carried out to compare the variance explained by internal vs. external factors.

Table 5. Backward regression model results for receptive grammar.

	B	SE	β	t	Sig	95% CIs	Part
(Constant)	-10.993	3.376		-3.256	.002	-17.736, -4.251	
HomMed	.326	.147	.193	2.215	.030	.032, .621	.187
SchInput	.044	.009	.448	5.122	.000	.027, .061	.432
EngUse	2.355	.506	.415	4.650	.000	1.343, 3.366	.392
AoO	.204	.056	.321	3.656	.001	.092, .315	.308
NonInte	1.231	.537	.197	2.291	.025	.158, 2.303	.193

Note. $R^2 = .538$, $F(5,65) = 15.111$, $P = .001$

Table 6. The comparison of R^2 s from CV and from multiple regression analysis.

English outcome	Significant predictors	R^2 s from CV	R^2 s from multiple regression analysis
EngPro	SchInput, HomMed, MotEng, AoO, ShoMem	21.2%	35.2%
EngCom	SchInput, HomMed, and EngUse	15.6%	27.5%
EngGra	SchInput, HomMed, and EngUse, AoO, and NonInte	43.8%	53.8%

Note. NB: AoO = age of onset in months; ShoMem = short-term memory score; NonInte = non-verbal IQ score; SchInput = total amount of English input; HomMed = weekly English input quality at home (the sources and frequency of using English media); MotEng = mothers' self-rated proficiency in English;

4.3 Regression analyses for receptive grammar skill

Backward regression was used to predict receptive grammar as well. As can be seen in the results shown in Table 5, the model was significant ($F(5,65) = 15.111$, $p < .000$) and accounted for 53.8% of the variance in receptive grammar skill. The predictors are nonverbal intelligence, age of onset, total amount of school input, home English media, and English usage. Of these, the total amount of school input was the best predictor (Beta = .448, partial $r = .536$).

In order to compare the impact from internal and external factors respectively, forced entry regression was used, where external factors (SchInput, HomMedia, EngUse) were entered as a first block, followed by internal factors (AoO and NonInte) as a second block. Subsequently, the R^2 change was calculated. The data show that external factors alone explained 40.7% of the variance in receptive grammar scores ($R^2 = .407$, $F(3,67) = 15.327$, $p < .001$), and just 13.1% of additional variance was explained by entering internal factors (R^2 changed = .131, $F(2,65) = 9.176$, $p < .001$). A reversed order was also conducted and the result was similar: only 8.7% of the variance was explained by internal factors ($R^2 = .087$, $F(2,68) = 3.237$, $p < .045$), while external factors explained as much as an additional 45.1% of the variance (R^2 change = .451, $F(3,65) = 21.112$, $p < .001$).

4.4 Leave-one-out cross-validation

In order to gain more information about the models' predictive ability for different data from the same

population, leave-one-out cross-validation (CV) was used on each of the three best fitting models of English outcomes. As expected, the R^2 s from CV were smaller than the R^2 s from the multiple regression analyses. The specific R^2 s from CV are listed in Table 6.

4.5 Bayes factor (BF)

The BFs generally supported the best fitting model selected by the backward regression approach. With respect to the productive vocabulary skill and receptive grammar skill, the best models selected by the two approaches are exactly the same. However, with respect to the receptive vocabulary skill, the model selected with BF contained one more predictor, "Age of Onset", than the best one selected by backward regression. The ratio of both models' BF values is about 1.01, indicating that both models are equally supported by the data. Thus, for the receptive vocabulary skills, we stayed with the backward regression model in Table 4.

5. Discussion

The current study investigated how internal factors and external factors influence the acquisition of English receptive vocabulary (PPVT-4), English productive vocabulary (EOWPV-2), and English receptive grammar (TROG-2) of very young Chinese learners of English in China. The present study is one of the few studies that examined both internal and external factors on various domains of the same population of very young learners of

English in the instructional setting (e.g., Unsworth et al., 2014).

The best fitting models of the three English skills contain similar external factors, but different internal factors. The total amount of school input and home English media environment were significant predictors for all of the three models. The other predictors were selected by different models: English usage for receptive vocabulary and grammar; mothers' English proficiency for productive vocabulary only; age of onset for productive vocabulary and receptive grammar; short-term memory for productive vocabulary only; and nonverbal intelligence for receptive grammar only. The results are generally in line with the previous child ESL and EFL studies (Unsworth et al., 2014).

Contrary to Paradis' findings (2011) on child ESL learners, in the current EFL study, external factors explained more variance in the three outcome variables than internal factors. Both input quantity (hours of school input) and input quality (number of different English media at home) were robust predictors regardless of the language domains. The different results from EFL and ESL settings are probably due to variation in access to English exposure. Compared to child ESL learners in general, who may frequently hear and use English at school, in the community, and through media, the exposure of child EFL learners to English is quite limited. The only comparatively stable English input in the current study was from the English school, where children could have English class for about 2 hours per week, with the exception of holidays and absences. The rest of English input and usage appeared to depend on the family, and parents varied substantially in the provision of English media and in the creation of opportunities for their children to use English. Therefore, the external factors could be more sensitive than internal factors in capturing differences in the outcome of children's English.

5.1 External factors

The total amount of school input, a more accurate measurement of cumulative English input quantity than length of exposure, was found to significantly predict all three aspects of English skills and was therefore the best predictor for receptive grammar. This is in line with previous findings of early ESL and EFL studies (Chondrogianni & Marinis, 2011; Muñoz, 2011, 2014; Unsworth et al. 2014). Despite the general agreement on the role of the amount of input in second or foreign language acquisition, "there is little consensus about which linguistic domains should be affected or to what extent" (Unsworth, 2013b, p. 86). Results of the current study indicate that input quantity might positively influence both vocabulary and grammar acquisition, and have more effect on the latter in an EFL setting.

In terms of input quality, the current paper used home English media to capture the frequency and variety of English input at home. This was found to be a significant predictor of all three English skills and was considered the best predictor for both productive and receptive vocabulary. This confirms previous findings on input quality in general (e.g., Jia & Fuse, 2007; Place & Hoff, 2011). Compared to ESL settings, child EFL learners rely more on media input since they barely have any native English-speaking friends and seldom use English at home and in the community. The current study demonstrated that English media can play an active role in children's English learning. According to the comments made by parents in the questionnaire, frequent and short English input, such as a cartoon clip or a song, seems to have attracted children's attention and maintained their motivation. Among the different kinds of media input, English movies and electronic devices were most frequently used. It is worth noting that several children in the current study had hours of movie input, but performed below average. This could be due to the fact that the level of the children's English proficiency and the input materials do not match. Children might be very interested in Disney movies at the beginning, but if the language is too complicated, or the episode is too long, they might get frustrated after five minutes of watching and gain little from such exposure. Electronic devices, such as tablet computers and smart-pen readers, were used as language toys by many parents in the current study and children in general liked the English games. Suggestions on how to use these devices and how to support children to better acquire English could be beneficial for parents. Currently, an organization that provides such a support is hard to find, as mentioned by most parents during the interview. The website for Dutch bilingual children (Brasileiro, Pinto & Unsworth, 2011) might serve as a good example for researchers of EFL learners to bring new findings into practice.

Children's own English use has been highlighted in several recent studies as a significant predictor of child ESL learners' language acquisition (Bohman et al. 2010; Montrul, 2008; Paradis, 2011) and a similar result was found by the current study. English use was a significant predictor for both receptive vocabulary and grammar. Bohman et al. (2010) argued that it might be more important for morphosyntax than vocabulary development, since language practice is more likely to facilitate the development of accuracy and automaticity with grammatical constructions than with words. Their hypothesis was based on the results of L2 production and the current study indicates that it might also apply to L2 reception, since English use explained more variance in receptive grammar skills than in receptive vocabulary.

Mothers' English proficiency was a significant predictor only of English productive vocabulary. This

might be due to the limited use of English and to the way the vocabulary was reviewed at home. Since most of the mothers in the current study have low oral English proficiency, they were wary of using English with their children. The situation in which English was most frequently used was when they helped their children review the words learned at Happy English. Mothers with higher English proficiency were more likely to practice words with their children and thus provided more chances for their children to use English words.

5.2 Internal factors

Age of onset emerged as a significant predictor of both productive vocabulary and receptive grammar. Older age of onset, predicting better learning outcome, indicates that children with greater cognitive and linguistic maturity were more advanced in their FL development. The cognitive maturity could be attributed to children's normal physical development, but it could also be related to the formal education that children take before the onset of English learning. In the current study, children's onset age of English learning is highly correlated with the length of kindergarten attendance before starting English education ($r = .89$). An additional round of backward regressions without AoO revealed that the length of formal education before AoO could significantly predict children's receptive grammar; however, the explained variance was smaller than in the models using AoO. Studies in the future should deconstruct the predictor AoO and explore its relationship with more environmental factors in the instructional setting (e.g., length of formal education before English learning). The current findings are in line with what has been found in child ESL studies (Golberg et al., 2008; Chondrogianni & Marinis, 2011) and child EFL studies (Muñoz, 2006; García Mayo, 2003; Cenoz, 2003; Ojima et al., 2011). At least in the short run, older FL learners were found to outperform their younger counterparts in general. It is difficult to draw a conclusion regarding the appropriate age to start learning a foreign language, however, passionate parents believing in "the younger, the better" should be informed that an early start does not necessarily guarantee a better ultimate achievement. External factors, especially the amount and quality of input should be paid special attention to.

Components of language aptitude were found to be a significant predictor for different language skills. Short-term memory, as measured by digit and non-word repetition subtests of the CTOPP and hand movement repetition subtest of the K-ABC, was a significant predictor of productive vocabulary, while analytical reasoning ability, as measured by nonverbal intelligence task of Raven's, significantly predicted English receptive

grammar. This is consistent with previous findings of child ESL and EFL studies (Alexiou, 2009; French & O'Brien, 2008; Harley & Hart, 1997; Masoura & Gathercole, 1999; Service, 1992, Paradis, 2011). However, the question remains: why was analytical reasoning ability more important than memory in predicting receptive grammar. Some child L2 studies found that short-term memory has a significant impact on both vocabulary and grammar acquisition (e.g., Harley & Hart, 1997; Paradis, 2011). A closer look at these studies revealed that children in this research were from naturalistic settings where more L2 input and practice were available than in the instructional settings. This implies that the memory-based approach of L2 learning might heavily rely on the L2 environment. When external factors are less favorable, as with the children in the current study where L2 input and output are scarce, the power of certain internal factors may be lessened. In these contexts, analytical reasoning ability might emerge as a more significant factor than memory in dealing with sentences, because it helps children to better organize the intensive and complicated information (Milton & Alexiou, 2006).

6. Limitations

The present study has several limitations. Firstly, this is a cross-sectional study and conclusions about causality could therefore not be generated. Later studies could adopt a longitudinal design to track trajectories of development over time. Secondly, teachers' English proficiency and classroom instructions should also be taken into consideration (Unsworth et al., 2014). Finally, the elicitation of home English input quantity could be improved. Future researchers could use the language diary approach (De Houwer & Bornstein, 2003) to track children's daily schedule for several weeks. These records would more precisely reflect children's language input in different situations and thus provide us more information about the impact of various inputs on language development.

7. Conclusion and implications

In this paper, we investigated the influence of internal and external factors on very young child EFL acquisition. The three aspects of language skills studied in this paper (English receptive vocabulary, English productive vocabulary and English receptive grammar) have similar external predictors but different internal predictors. Input quantity and quality, measured as the total amount of school input and the number of home English media, are significant predictors for all three aspects. English use also plays an important role in the receptive vocabulary and grammar. Mothers' oral English proficiency significantly predicted only productive vocabulary, probably due

to limited use of English and the way vocabulary was reviewed at home. In terms of internal factors, children with older onset age outperformed those with younger onset age on productive vocabulary and receptive grammar. Moreover, better short-term memory was found to facilitate productive vocabulary acquisition and better analytical reasoning ability might promote receptive grammar learning. This is generally in line with the previous findings.

Contrary to Paradis' (2011) study on child ESL learners in the naturalistic setting, external factors explained more variance than internal factors in the current child EFL study. This discrepancy might result from the different input environment, wherein children from an EFL setting usually have less English exposure than peers in an ESL setting. The comparison of the results from the current study, and from Paradis (2011), indicates that external factors are more context-sensitive and different populations might therefore vary from each other in that respect. While internal factors make a stable and important contribution to child L2 acquisition, external factors can starkly influence the outcome on top of that, even to the extent that they explain more of the outcomes than the internal factors. Both internal and external factors play a crucial role. Therefore, our next priority could be the exploration of how better to use external factors, since they could be manipulated and can drastically boost the learning outcomes.

The children in the current study are part of the millions of young English learners in China. Against the backdrop of globalisation, more and more parents consider English to be an important part of their children's education and start to send them to various early English training programs at a very young age. Due to the restrictions of the language environment in China, children's English input and use are quite limited. At school, English classes are only held for one or two hours per week and, at home, parents are reluctant to use English with their children for fear of negatively influencing their child's proficiency. This situation is quite common in East Asia, especially in Japan, South Korea, and China (Butler, 2013). Various kinds of home English materials could have the potential

to ameliorate the EFL environment by providing more English input and practice opportunities.

Rhymes, songs, stories, and activities with a ritualistic character have been found to provide input for both lexical and grammatical learning (Sokolov & Snow, 1994). Moreover, different media devices, such as computers and tablets, deliver these contents to child foreign language learners with interesting and engaging experiences. L2 input from such media devices might significantly boost children's learning motivation. As a result, future research should explore in detail the approach of using such media. Insights from early L1 and child L2 studies on using media (e.g., Bus, Verhallen & de Jong, 2009) could be introduced and verified for child FL learners in an instructional setting.

Appendix I. Language Exposure Questionnaire

Part 1 Background

Child's name	Test number	Gender
Place of birth	Date of birth	Date of testing
Mother's name	Mother's occupation	Mother's highest education
Father's name	Father's occupation	Father's highest education
Sibling's name	Sibling's age	

Part 2 English learning and usage

1. From whom and where did your child first come into contact with English?
2. From when did your child receive consistent and significant exposure to English?
3. Think about the English exposure of your kid each normal week at home (and at kindergarten if applicable)

		English TV			English electronic devices (e.g., smart-pen reader)	English books	English games	Kindergarten/ School
Weekly exposure (in minutes)	English TV programs made in China	English programs made in speaking countries	English videos (e.g., movie “Three pigs”)	English audios (e.g., Action Verses)				
Mon.								
Tue.								
Wed.								
Thur.								
Fri.								
Sat.								
Sun.								
Total (in minutes)								
Materials								

4. How long does your child use English with you in a normal week (in minutes) at home?
Mother: ____ Father: ____ Other relatives ____
5. Think about the oral English proficiency of the following members at home
Mother: (Poor 1 – Fair 2 – Average 3 – Good 4 – Excellen5)
Father: (Poor 1 – Fair 2 – Average 3 – Good 4 – Excellen5)
Other relatives: (Poor 1 – Fair 2 – Average 3 – Good 4 – Excellen5)
6. Has any Chinese teacher taught your child English besides Happy English and kindergarten/ school? If so, how long was it? (in minutes)
7. Has any native English speaker taught your child English besides Happy English and kindergarten/ school? If so, how long was it? (in minutes)
8. How many times does your child review English words learnt at Happy English in general?
9. Where does your child use English regularly?
 - a. Happy English
 - b. at home
 - c. on vacation
 - d. At kindergarten/school
 - e. English stage performance
 - f. Other places, such as ____
10. What’s your biggest concern so far to help your child learn English?
 - a. Lack of appropriate materials
 - b. Lack of good teaching approach and skills
 - c. Lack of confidence in my oral English
 - d. Lack of time due to being busy
 - e. I don’t want to give my child too much burden
 - f. I believe that good English acquisition relies on language aptitude
 - g. I believe that good English acquisition relies on English teachers
 - h. If you have other concerns, please write them down: _____

Appendix II English exposure at kindergarten/school

1. When did your child start to learn English at kindergarten/school?
2. How long does he/she learn English there?
3. How many classes does he/she have per week and how long does each class last?
4. Did he/she have a Chinese or an English teacher? What proficiency level does the teacher have?

References

Alexiou, T. (2009). Young learners’ cognitive skills and their role in foreign language vocabulary learning. In M. Nikolov (Ed.), *Early learning of modern foreign languages: Processes and outcomes* (pp. 46–61). Clevedon: Multilingual Matters.

- Asher, J. (1996). *Learning another language through actions* (5th ed.). Los Gatos, CA: Sky Oaks Productions.
- Bishop, D. (2003). *Test for reception of grammar* (2nd ed.). London: Pearson Assessment.
- Blom, E., Paradis, J., & Sorenson Duncan, T. (2010). The acquisition of 3SG-s by L2 children: Domain-general or domain-specific learning? *Poster Presented at the Child Language Seminar*, City University, London, UK.
- Blom, E., & Paradis, J. (2014). Sources of individual differences in the acquisition of tense inflection by English second language learners with and without specific language impairment. *Applied Psycholinguistics, First View*, 1–24. doi:10.1017/S014271641300057X
- Bohman, T. M., Bedore, L. M., Pena, E. D., Mendez-Perez, A., & Gillam, R. B. (2010). What you hear and what you say: Language performance in Spanish-English bilinguals. *International Journal of Bilingual Education and Bilingualism*, 13(3), 325–344. doi:10.1080/13670050903342019
- Brasileiro, I., Unsworth, S. & Pinto, M. Growing up bilingual. Retrieved from <http://www.tweetaligoprocieen.org/author/dafne/>
- Brownell, R. (2010). *Expressive one-word picture vocabulary test*. San Antonio: Pearson.
- Bus, A. G., Verhallen, M. J. A. J., & de Jong, M. T. (2009). How onscreen storybooks contribute to early literacy. In A. G. Bus, & S. B. Neuman (Eds.), *Multimedia and literacy development: Improving achievement for young learners* (pp. 153–167). New York: Routledge Education: Taylor & Francis Group.
- Butler, Y. G. (2013). Parental factors and early English education as a foreign language: A case study in mainland china. *Research Papers in Education*. doi:10.1080/02671522.2013.776625.2013.
- Carroll, J. B., & Sapon, S. (2002). *Modern Language Aptitude Test: Manual 2002 Edition*. N. Bethesda, MD: Second Language Testing, Inc.
- Cenoz, J. (2003). The influence of age on the acquisition of English: General proficiency, attitudes and code-mixing. In M. García Mayo, & M. García Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 77–93). Clevedon: Multilingual Matters.
- 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(3), 223–248.
- Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York: Routledge.
- Cummins, J. (1984). *Bilingualism and special education*. Clevedon, UK: Multilingual Matters.
- De Houwer, A., & Bornstein, M. (2003). Balancing on the tightrope: Language use patterns in bilingual families with young children. *Presented at the Fourth International Symposium on Bilingualism*, Tempe, Arizona, April 30 to May 3.
- Dunn, L., & Dunn, L. (2007). *Peabody picture vocabulary test* (4th ed.). Circle Pines, MN: American Guidance Service.
- French, L. M., & O'Brien, I. (2008). Phonological memory and children's second language grammar learning. *Applied Psycholinguistics*, 29(3), 463–487. doi:10.1017/S0142716408080211
- García Mayo, M. P. (2003). Age, length of exposure and grammaticality judgments in the acquisition of English as a foreign language. In M. P. García Mayo, & M. L. García Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 94–114). Clevedon, Avon: Multilingual Matters.
- García Mayo, M. P., & García Lecumberri, M. L. (Eds.). (2003). *Age and the acquisition of English as a foreign language: Theoretical issues and field work*. Clevedon: Multilingual Matters.
- Genesee, F., & Hamayan, E. (1980). Individual differences in second language learning. *Applied Psycholinguistics*, 1(01), 95–110. doi:10.1017/S0142716400000758
- Golberg, H., Paradis, J., & Crago, M. (2008). Lexical acquisition over time in minority first language children learning English as a second language. *Applied Psycholinguistics*, 29(1), 41–65. doi:10.1017/S014271640808003X
- Gutierrez-Clellen, V. F., & Kreiter, J. (2003). Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics*, 24(2), 267–288. doi:10.1017/S0142716403000158
- Harley, B. (1998). The outcomes of early and later language learning. In M. Med (Ed.), *Critical issues in early second language learning* (pp. 26–31). Scott Foresman-Addison Wesley.
- Harley, B., & Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition*, 19(03), 379–400.
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, 26(1), 55–88. doi:10.1016/j.dr.2005.11.002
- Hyltenstam, K., & Abrahamsson, N. (2003). Maturation constraints in SLA. In C. Doughty, & M. Long (Eds.), *The handbook of second language acquisition* (Oxford: Blackwell. ed., pp. 540–558). Oxford: Blackwell.
- Sokolov, J. L., & Snow, C. E. (1994). The changing role of negative evidence in theories of language development. In C. Gallaway and B. J. Richards. (Ed.), *Input and interaction in language acquisition*, (pp.38–55). Cambridge: Cambridge University Press.
- Jia, G., & Aaronson, D. (2003). A longitudinal study of Chinese children and adolescents learning English in the united states. *Applied Psycholinguistics*, 24(1), 131–161. doi:10.1017/S0142716403000079
- 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. doi:10.1044/1092-4388(2007/090)
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430), 773–795. doi:10.1080/01621459.1995.10476572
- Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman assessment battery for children*. Pearson.

- Kersten, K., Rohde, A., Schelletter, C., Steinlen, A. (2010). *Bilingual Preschools. Volume I: Learning and Development*. Trier: WVT.
- Masoura, E. V., & Gathercole, S. E. (1999). Phonological short-term memory and foreign language learning. *International Journal of Psychology, 34*(5–6), 383–388.
- Milton, J., & Alexiou, T. (2006). Language aptitude development in young learners. In C. Abello-Contesse, R. Chacón-Beltrán, M. D. López-Chiménez & M. M. Torreblanca-López (Eds.), *Age in L2 acquisition and teaching* (pp. 177–192). Oxford: Peter Lang.
- Ministry of Education of the People's Republic of China. (2001). *Quanrizhi yiwu jiayu putong gaoji shongxue yingyu kecheng biaoqun [english curriculum standards for full-time compulsory education and senior secondary schools]*, Beijing: Beijing Normal University Press.U.S.C.
- Montrul, S. (2008). *Incomplete acquisition in bilingualism: Re-examining the age factor*. Amsterdam: John Benjamins.
- Muñoz, C. (2006). The BAF project: Research on the effects of age on foreign language acquisition. In C. Abello-Contesse, R. Chacón-Beltrán, M. D. López-Jiménez & M. M. Torreblanca-López (Eds.), *Age in L2 acquisition and teaching* (pp. 81–92). Bern: Peter La.
- Muñoz, C. (2008). Symmetries and asymmetries of age effects in naturalistic and instructed L2 learning. *Applied Linguistics, 29*(4), 578–596. doi:10.1093/applin/amm056
- Muñoz, C. (2011). Input and long-term effects of starting age in foreign language learning. *International Review of Applied Linguistics in Language Teaching, 49*(2), 113–133. doi:10.1515/iral.2011.006
- Muñoz, C. (2014). Contrasting effects of starting age and input on the oral performance of foreign language learners. *Applied Linguistics, 35*(4), 463–482. doi:10.1093/applin/amu024
- Nikolov, M., & Mihaljević Djigunovic, J. M. (2011). All shades of every color: An overview of early teaching and learning of foreign languages. *Annual Review of Applied Linguistics, 31*, 95–119. doi:10.1017/S0267190511000183
- Ojima, S., Matsuba-Kurita, H., Nakamura, N., Hoshino, T., & Hagiwara, H. (2011). Age and amount of exposure to a foreign language during childhood: Behavioral and ERP data on the semantic comprehension of spoken English by Japanese children. *Neuroscience Research, 70*(2), 197–205. doi:10.1016/j.neures.2011.01.018
- Oller, D. K., & Eilers, R. E. (2002). *Language and literacy in bilingual children*. Clevedon, UK: Multilingual Matters.
- Paradis, J. (2009). *Oral language development in French and English and the role of home input factors*. (Report for the Conseil scolaire centre-nord[north-central school board]). Edmonton, Alberta.
- Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism, 1*, 213–237.
- Paradis, J., Nicoladis, E., Crago, M., & Genesee, F. (2011). Bilingual children's acquisition of the past tense: A usage-based approach. *Journal of Child Language, 38*(3), 554–578. doi:10.1017/S0305000910000218
- Place, S., & Hoff, E. (2011). Properties of dual language exposure that influence 2-year-olds' bilingual proficiency. *Child Development, 82*(6), 1834–1849. doi:10.1111/j.1467-8624.2011.01660.x
- Raven, J. C., Court, J. H., & Raven, J. (1995). *Coloured progressive matrices manual*. London: Lewis.
- Refaeilzadeh, P., Tang, L., & Liu, H. (2009). Cross-validation. *Encyclopedia of Database Systems*, 532–538.
- Sawyer, M., & Ranta, L. (2002). Aptitude, individual differences, and instructional design. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 319–353). Cambridge, UK: Cambridge University Press.
- Scheele, A. F., Leseman, P. P. M., & Mayo, A. Y. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics, 31*(1), 117–140. doi:10.1017/S0142716409990191
- Service, E. (1992). Phonology, working memory, and foreign-language learning. *Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology, 45*(1), 21–50.
- Skehan, P. (1986). Cluster analysis and the identification of learner types. In V. Cook (Ed.), *Experimental approaches to second language learning* (pp. 81–94). Oxford: Pergamon Press.
- Stevens, G. (2006). The age-length-onset problem in research on second language acquisition among immigrants. *Language Learning, 56*(4), 671–692. doi:10.1111/j.1467-9922.2006.00392.x
- Sun, H., de Bot, K., & Steinkrauss, R. (In press). A multiple case study on the effects of temperamental traits in Chinese preschoolers learning English. *International Journal of Bilingualism*. Advance online publication doi:http://ijb.sagepub.com/content/early/2014/06/04/1367006914534332.full.pdf?ijkey=1qHVUeTWzbaQoZn&keytype=ref.
- Thorn, A. S. C., & Gathercole, S. E. (1999). Language-specific knowledge and short-term memory in bilingual and non-bilingual children. *Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology, 52*(2), 303–324. doi:10.1080/027249899391089
- Unsworth, S. (2005). *Child L2, adult L2, child L1: Differences and similarities. A study on the acquisition of direct object scrambling in Dutch*. (Unpublished Phd thesis). Utrecht University, the Netherlands.,
- Unsworth, S., Hulk, A., & Marinis, T. (2011). Internal and external factors in child second language acquisition. *Introduction to Spotlight Issue of Linguistic Approaches to Bilingualism, 1*, 207–212.
- Unsworth, S., Persson, L., Prins, T., & de Bot, K. (2014). An investigation of factors affecting early foreign language learning in the Netherlands. *Applied Linguistics*,
- Unsworth, S. (2013a). Assessing the role of current and CUMULATIVE exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism-Language and Cognition, 16*(1), 86–110. doi:10.1017/S1366728912000284

- Unsworth, S. (2013b). Current issues in multilingual first language acquisition. *Annual Review of Applied Linguistics*, 33, 21–50. doi:[10.1017/S0267190513000044](https://doi.org/10.1017/S0267190513000044)
- Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L-1/L-2 acquisition and frequency of input. *Applied Psycholinguistics*, 22(2), 217–234. doi:[10.1017/S0142716401002041](https://doi.org/10.1017/S0142716401002041)
- Wagner, R., Torgesen, J., & Rashotte, C. (1999). *Comprehensive test of phonological processing*. Austin, TX: Pro-Ed.
- Yin, P., & Fan, X. (2001). Estimating R2 shrinkage in multiple regression: A comparison of different analytical methods. *The Journal of Experimental Education*, 69(2), 203–224.