Longitudinal predictors of listening comprehension in bilingual primary school-aged children

LANGUAGE LEARNING
A Journal of Research in Language Studies
Language Learning ISSN 0023-8333

EMPIRICAL STUDY

Longitudinal Predictors of Listening Comprehension in Bilingual Primary School-Aged Children

Alessandra Valentini[a,b,c] and Ludovica Serratrice[a,d]

[a]School of Psychology and Clinical Language Sciences, University of Reading, [b]School of Psychology, University of Surrey, [c]School of Human Sciences, University of Greenwich, and [d]Department of Language and Culture, UiT the Arctic University of Norway

Abstract: Research on monolingual children has shown that listening comprehension is predicted by a range of language and cognitive skills; less is known about predictors of listening comprehension in bilingual children and about the role of language input. This study presents longitudinal data on predictors of English listening comprehension in 100 bilingual children between the ages of 5;8 and 6;8 years. The children were tested three times on their literal and inferential comprehension of stories. Vocabulary, morphosyntax, attention, and memory were included as predictors of listening comprehension alongside a measure of English input. The children showed growth over time in both literal questions and global inference questions, with performance on local inferences remaining stable over time. Vocabulary depth and morphological knowledge explained listening comprehension abilities in all types of questions, but not their growth; that is, all children improved in comprehension over time regardless of their

---

Alessandra Valentini and Ludovica Serratrice were jointly responsible for the conception and the design of the study, for the interpretation of the data, and for drafting and revising the manuscript. Alessandra Valentini was solely responsible for the data collection and for the data analysis. This research was conducted while Alessandra Valentini was affiliated with the University of Reading and written while she was affiliated with the University of Surrey. The authors have no conflict of interest to declare.

Correspondence concerning this article should be addressed to Alessandra Valentini, School of Human Sciences, University of Greenwich, Dreadnought Building, 30 Park Row, Greenwich, London, SE10 9LS, United Kingdom. Email: a.valentini@greenwich.ac.uk

The handling editor for this article was Theres Grüter.

DOI: 10.1111/lang.12513

initial morphological and vocabulary depth skills. English input had a mediated effect on listening comprehension via morphological knowledge and vocabulary depth, but no direct effect.

Keywords listening comprehension; bilingualism; language acquisition

Introduction

Listening comprehension, that is, the ability to understand spoken language, is essential for successful communication and has a prominent role in literacy acquisition and reading comprehension (Hoover & Gough, 1990). Understanding spoken language relies on the ability to retain and store information to create an integrated mental model of the state of affairs, that is, a situation model (Kintsch, 1994) including information that is mentioned overtly and information that is only suggested by the text. In making inferences, listeners go beyond what is stated explicitly and make informed guesses about what is implicitly intended. Inferencing skills are thus necessary to make connections between pieces of information in the text (local inferences), or with preexisting background knowledge outside of the text (global inferences). Local inferences are necessary for integrating two propositions through the mapping of related words, for example, between synonyms or category-exemplar pairings or for resolving anaphoric dependencies. Global inferences, on the other hand, are connections made between information in the text and general background knowledge acquired previously, for example, through personal experiences or reading.

A growing body of research has investigated how monolingual children use different language and cognitive skills in listening comprehension (Alonzo, Yeomans-Maldonado, Murphy, Bevens, & Language and Reading Research Consortium [LARRC], 2016; Currie & Cain, 2015; de Bree & Zee, 2020; Florit, Roch, & Levorato, 2011, 2013; Kim, 2016; Lepola, Lynch, Laakkonen, Silvén, & Niemi, 2012; Strasser & Del Rio, 2014). However, research has not yet established how linguistic and cognitive skills predict listening comprehension concurrently and longitudinally in children who speak more than one language. Because of the distributed nature of exposure to their languages, bilingual children also offer a unique opportunity for investigating the role of relative amount of input in the process of listening comprehension in the corresponding language.

Our study had three distinct aims. First, we intended to explore the role of foundational cognitive skills (attention and memory), foundational language skills (vocabulary and grammar), and higher order skills (comprehension monitoring and inferencing) in predicting listening comprehension abilities and their growth in bilingual children. Second, we wanted to investigate potential differences between literal and inferential comprehension in terms of their growth and their predictors. Third, we modeled the role that the amount of input in the language of schooling (English) plays in listening comprehension, considering its possible direct effect as well as its indirect effect through other language skills.

Background Literature

Predictors of Listening Comprehension

Predictors of listening comprehension can be broadly grouped into three categories: foundational cognitive skills, foundational language skills, and higher order skills (Kim, 2016).

The foundational cognitive skills most often associated with listening comprehension are memory, attention, and IQ (Kim, 2016; Strasser & Del Rio, 2014). Most of the research on memory and listening comprehension has focused on the effects of short-term memory, that is, the capacity of the short-term storage, and of working memory (Baddeley, 1986), that is, the ability to manipulate information from short-term memory (Florit, Roch, Altoé, & Levorato, 2009; LARRC, Jiang, & Farquharson, 2018). Results have been mixed, with some studies showing the effects of measures of both short-term and working memory (Florit et al., 2009, 2013), although other studies have found only an effect of working memory (Silva & Cain, 2015). Given the mixed results, we considered measures of both types of memory in our research.

Another important cognitive skill related to comprehension is attention, specifically the ability to focus on the task of listening long enough to process the information and the ability to inhibit irrelevant stimuli. However, although attention skills have been associated with language comprehension (LARRC et al., 2018), some studies have suggested that attention, measured using behavioral checklists, might have only an indirect effect on listening comprehension via its effect on other language skills (Kim, 2016). In our study, we focused specifically on the role of auditory attention, measured directly rather than using a checklist.

When foundational language skills have been considered, vocabulary and morphosyntax have repeatedly been associated with language comprehension (Alonzo et al. 2016; Kim, 2016; Strasser & Del Rio, 2014); understanding single words and their structural relationship within a sentence are the essential first steps to comprehending the meaning of a text. When considering vocabulary, recent research has highlighted the importance of distinguishing between breadth and depth of vocabulary knowledge (Ouellette, 2006) and the possibility that bilingual children might lag behind monolingual peers in one of these aspects of vocabulary but not the other (Dixon, Thomson, Fricke, 2020). Vocabulary breadth, defined as the number of entries in the mental lexicon, has been the focus of research in both reading (Eason, Goldberg, Young, Geist, & Cutting, 2012; Muter, Hulme, Snowling, & Stevenson, 2004) and listening comprehension (Silva & Cain, 2015). More recent research has suggested that vocabulary depth, namely the extent of word-related knowledge and the density of a speaker's semantic network (i.e., the number of links between words), plays a crucial role in comprehension (de Bree & Zee, 2020; Lepola et al. 2012). In essence, the quality of lexical knowledge, which was operationalized as vocabulary depth, predicted comprehension in these studies (Perfetti, 2007). Studies that have used both vocabulary breadth and depth tend to report similar (Florit et al., 2009, 2013) or stronger effects of vocabulary depth on comprehension (Strasser & Del Rio, 2014). When differentiating vocabulary breadth and depth, it is further important to distinguish between tasks that tap into expressive or receptive vocabulary. To avoid the confound of task difficulty, in our study, we used several different measures of vocabulary depth, none of which required the production of definitions because this can be particularly difficult for bilingual children who often have a gap between receptive and expressive skills (Gibson, Oller, Jarmulowicz, & Ethington, 2012).

Morphosyntactic knowledge, namely children's ability to comprehend and produce syntactic constructions (e.g., passives vs. actives; subordinates vs. main clauses) and inflectional and derivational morphology (e.g., suffixes for tense changes, or changes to part of speech), has previously been linked to listening comprehension. For example, Kim (2016) found a direct effect of syntactic awareness, that is, the ability to distinguish between grammatical and nongrammatical constructions, on listening comprehension in 6- to 7-year-olds, as well as an effect on listening comprehension mediated via comprehension monitoring. In Babayiğit (2014), morphosyntactic skills measured via sentence repetition were significant predictors of listening comprehension for 9- to 10-year-olds. However, not all studies have found an effect of morphosyntactic knowledge on listening comprehension (e.g., see Alonzo et al., 2016). Overall, although it would be logical to expect a positive relationship between morphosyntactic knowledge and listening comprehension, this relationship may vary as a function of population and assessment method. In this study we included separate measures of syntactic and morphological knowledge to tap into these two relatively separate constructs.

Among higher order cognitive skills, comprehension monitoring and inferencing have been the most widely studied. Comprehension monitoring is the ability to check one's own understanding of a text and the ability to detect any inconsistencies within the text itself (Ruffman, 1996). Several studies have shown a significant effect of comprehension monitoring on listening comprehension (Kim, 2016) in children as young as 5 years of age (Strasser & Del Rio, 2014).

Making inferences, that is, the ability to link information within a text (local inferences) or outside of a text (global inferences), is part of the comprehension process itself. However, several studies have considered inference making as a predictor of broader comprehension skills, finding a link between inferencing and listening comprehension (Florit et al., 2011; Kim, 2016). Most studies have employed verbal inferencing tasks where children listen to or read passages and answer inferential questions, but a few studies used wordless picture books to measure inferencing without relying on children's linguistic abilities. These studies observed a correlation between these wordless picture tasks and reading and listening comprehension (Lepola et al., 2012; Paris & Paris, 2003). Similarly, in our study, we used a wordless picture task to explore the effect of children's inference-making abilities independently of their verbal language abilities.

Aspects of Listening Comprehension

Literal understanding (what is explicitly said in a text) and the ability to make local and global inferences underpin listening comprehension. However, the predictive power of foundational cognitive and language skills may vary depending on the aspect of comprehension investigated. Literal comprehension requires memory for the details of a story, but local inferences draw on syntactic, semantic, and discourse knowledge, and global inferences rely on comprehenders' semantic knowledge as well as their general world knowledge. Because of these different demands, children tend to find literal questions easier than inferential questions (Eason et al., 2012), and inferential questions that rely on background knowledge are generally harder than those requiring text connections (Barnes, Dennis, & Haefele-Kalvaitis, 1996).

As for predictors of different aspects of comprehension, a few studies have highlighted the importance of vocabulary depth for global inferences rather than for literal comprehension (Cain & Oakhill, 2014; Currie & Cain, 2015). Other studies have shown a reciprocal relationship between vocabulary and the ability to answer inferential questions (LARRC, Currie, & Muijselaar, 2019). Working memory has also been highlighted as a better predictor of the ability to answer inferential than literal questions (Alptekin & Erçetin, 2011). In our study, we differentiated between predictors of literal comprehension and predictors of the ability to make local and global inferences.

Effect of Language Input on Bilingual Children's Comprehension

Most studies of predictors of listening comprehension have focused on monolingual participants. When considering bilingual children, several studies have explored predictors of reading comprehension (Bowyer-Crane, Fricke, Schaefer, Lervåg, & Hulme, 2017; Melby-Lervåg & Lervåg, 2014), but there has been little research on predictors of listening comprehension in the population of bilingual children. The few studies that explored this topic showed the importance of vocabulary for listening comprehension (Burgoyne, Kelly, Whiteley, & Spooner, 2009; Hutchinson, Whiteley, Smith, & Connors, 2003). Babayiğit and colleagues (Babayiğit, 2014; Babayiğit & Shapiro, 2020) showed similar levels of predictions of vocabulary (breadth) and grammar measures (sentence repetition and syntactic knowledge) on listening comprehension in their monolingual and bilingual groups aged 9 to 10 years, and, in one study, they found an effect of age for the monolingual but not the bilingual group (Babayiğit & Shapiro, 2020). Age could be a proxy for maturation or for amount of language experience, especially for monolingual children, because older children have been exposed to language longer. The absence of an age effect for bilingual children could suggest that the amount of input, rather than age, is a better predictor in bilingual children because the amount of language experience is not just a function of age in this group. Hammer et al. (2012) did show that the amount of English exposure significantly predicted Spanish-English bilingual 5-year-olds' ability to retell a story. However, most studies on listening or reading comprehension in bilingual children have not explicitly modeled the predictive role of input, even when they have documented the amount of language exposure in their sample (Babayiğit, 2014; Babayiğit & Shapiro, 2020; Bowyer-Crane et al., 2017).

Language input is one of the strongest predictors of the rate of language development in monolingual and bilingual children (Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Pearson, 2007); measures of input explain variation in vocabulary development in bilinguals (Blom, 2010; Paradis, 2011, Sun, Yin, Amsah, & O'Brien, 2018) as well as variation in grammar knowledge and its development in both monolinguals (Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002) and bilinguals (Grüter & Paradis, 2014; Hoff, Quinn, & Giguere, 2018; Thordardottir, 2019). The study of the effects of input in bilingual acquisition is of particular theoretical and practical relevance.

Theoretically, individual differences in the amount of language experience in each language makes bilingualism an ideal test case to investigate how much input affects different aspects of language knowledge and their development. Practically, a deeper understanding of the relationship between input and language skills and their growth would allow practitioners to contextualize expectations for the achievement of bilingual children based on their language background. For these reasons, we included a measure of input in our analyses. Our aim was to explore both the direct effect of English input on listening comprehension as well as its indirect effect via other language skills like vocabulary and grammar.

The Present Study

Research Questions

In the light of the literature on listening comprehension in monolingual children and the relative paucity of research on bilingual children, we investigated the following questions:

1. How do foundational cognitive and language skills predict listening comprehension in bilingual children over time?

On the basis of previous research (Kim, 2016) our hypothesis was that grammar (either syntactic and/or morphological knowledge) and inferential skills would have a direct effect on listening comprehension, over and above other cognitive and language skills. We also expected possible direct effects of vocabulary (either breadth or depth; de Bree & Zee, 2020; Silva & Cain, 2015) and memory (Kim, 2016).

1. Is the growth of comprehension of literal information and of local and global inferences differentially affected by foundational cognitive and language skills?

Our hypothesis was that different aspects of listening comprehension might grow differently over time and that different predictors might explain different comprehension abilities. We specifically expected that vocabulary depth (Cain & Oakhill, 2014; Currie & Cain, 2015) would be a better predictor of local and global inferences than of literal comprehension.

1. How does amount of English input influence listening comprehension in bilingual children?

Our hypothesis was that English input would affect listening comprehension, but that this effect might be at least partly mediated by a direct relationship with vocabulary (Blom, 2010; Paradis, 2011) and grammar (Hoff et al., 2018; Huttenlocher et al., 2002).

Methods

Data, scripts, and supplementary materials (Valentini & Serratrice, 2022a, 2022b, 2022c) are available at IRIS (www.iris-database.org) and OSF (https://osf.io/2fa3c/?view_only=cc167f7a3484432b9d5d56e4b87d4032).

Participants

After receiving ethical approval by the research ethics committee at our institution, we recruited 100 bilingual children (48 girls; first testing session, M_{age} = 5;8 years, $SD = 0.29$) from nine schools in South East England; 89 participants completed all testing sessions. We included the data from all the children in the analyses whether or not they had participated in all sessions. We invited the children classified by their schools as having English as an additional language but not otherwise included in the schools' special education needs register to take part in the study. The Department for Education for England defines children with English as an additional language as those children "who were exposed to more than one language (which may include English) during early development" (Department for Education, 2020). We decided to include bilingual children according to this broad definition to reflect the linguistic diversity of bi/multilingual children in English primary schools. The children spoke 28 different languages (43% Polish, 7% Hindi, 5% Arabic, 5% French, 5% Romanian, 4% Malayalam, 4% Nepali) and were a representative sample of the current composition of multilingual classrooms in England.

Procedure

Parents and children gave their consent to participate in the study. We carried out the testing sessions at three time points between Year 1 and Year 2 of primary school: autumn–spring term (October–February) of Year 1 (Time 1), spring–summer term (April–July) of Year 1 (Time 2), and autumn–spring term (October–February) of Year 2 (Time 3). The children completed the tasks over two testing sessions at each time point to avoid fatigue. All the children completed the tasks in the same given order and in the same testing session (first or second session) at all time points.

The data reported here were part of a larger longitudinal study in which we collected several language and cognitive measures over time administered exclusively in English. Here we have reported the data from our specially designed listening comprehension task that the children completed at all time points, and Time 1 data for the language and cognitive tasks. This allowed us to investigate which language and cognitive measures at the start of the study would predict listening comprehension development over time.

Materials

Most of the measures used in our study were tasks from standardized tests used in previous studies with monolingual children. We used only raw scores in the analyses because these tests were normed on a monolingual population.

Foundational Cognitive Skills

We measured the children's general nonverbal abilities using the Matrix Reasoning subtest from the Wechsler Preschool and Primary Scale of Intelligence (Wechsler, 2013), where children had to choose the best picture to complete a pattern.

We assessed the children's attention skills using two subtests of the Test of Everyday Attention for Children (Manly, Robertson, Anderson, & Nimmo-Smith, 1998). The Score! subtest asks children to keep count of a series of randomly spaced sounds. The Walk Don't Walk subtest asks children to respond differently to different sounds.

We measured short-term memory skills using the forward digit span from the Clinical Evaluation of Language Fundamentals for the United Kingdom (CELF–4 UK; Semel, Wiig, & Secord, 2006); we assessed working-memory skills using the backward digit recall task from the CELF–4 UK, and a specially designed backward words recall task where the children were asked to repeat a series of words backward (see Appendix S2 in the online Supporting Information). For this task, we followed the testing procedure described in Florit et al. (2013).

Foundational Language Skills

To measure vocabulary breadth, we used the British Picture Vocabulary Scale (BPVS –3, Dunn, Dunn, & NFER, 2009), a word-picture matching task. To assess vocabulary depth, we administered three measures. The Synonyms and the Opposites subtests from the Test of Word Knowledge (Wiig & Secord, 1992) requires children to select the correct synonym or antonym for a given word. The experimenter presented words to the children in writing, and read them aloud. We further administered the Word Classes 1 subtest of the CELF–4 UK (Semel et al., 2006), where children have to identify conceptually related pictures and verbalize this relationship. This task yielded a receptive score (number of pairs correctly identified) and an expressive score (number of relationships correctly explained).

For morphosyntactic knowledge, the children completed two tasks: the Word Structure subtest from the CELF–4 UK (Semel et al., 2006), which requires the production of the final word in a given sentence with its correct morphological ending, and the short version of the Test of Reception of Grammar (Bishop, 2003; see Whiteside & Norbury, 2017, for the short version of this test), a sentence-picture matching task.

Higher Level Skills

We adapted the nonverbal inferencing task from Story A of the retelling task from the Multilingual Assessment Instrument for Narratives (Gagarina et al., 2012). The experimenter pointed at each picture in turn, describing its content but not giving any information to explain the pictures' meaning. After the picture presentation, the experimenter asked the children the 10 comprehension questions of the Multilingual Assessment Instrument for Narratives. This type of assessment of inferential abilities is similar to the procedure used in previous studies (Lepola et al., 2012; Paris & Paris, 2003).

The comprehension monitoring task followed a testing procedure supported by the literature (Ruffman, 1996) and required the children to indicate whether each of 12 short stories "made sense or not." The task was presented as a computer game, and the children indicated their choice by pressing the relevant button on the keyboard. Six of the stories presented a logical inconsistency whereby the third sentence contradicted a feature established by the first sentence (see Appendix S3 in the online Supporting Information). The children received one point for each story correctly categorized for a maximum score of 12; story order was randomized. This task had a level of reliability (Cronbach's $\alpha = .49$) in line with previous research (Cain, Oakhill, & Bryant, 2004).

Listening Comprehension

To assess listening comprehension we used a specially designed measure of listening comprehension. We administered the Understanding of Spoken Paragraphs subtest from the CELF–4 UK (Semel et al., 2006) at Time 1 and Time 3 to validate this measure.

This specially designed comprehension task consisted of three stories at each time point and measured the children's comprehension of literal information and their ability to make local and global inferences. Following Freed and Cain (2017), we attempted to alleviate the memory burden for our participants by dividing the stories into three parts with a mean length of 35 words each; the children listened to each part of the story through headphones, and the experimenter asked them two or three questions about what they had just heard (see Appendix S5 in the online Supporting Information). The children gave their answers orally, and their answers were recorded verbatim on an answer sheet. For each story, each child answered two literal and six inferential questions. The inferential questions were divided into three local inferences, where the children were required to make specific anaphoric inferences by connecting different parts of the text, and three global inferences, where the children had to integrate the story with their background knowledge. We asked four independent assessors with PhDs in linguistics to categorize each question as literal or as requiring a local or a global inference. We further assessed any question with less than 100% agreement and replaced that question if it was problematic.

To avoid effects of repeated exposure, we devised nine different stories for this specially designed listening comprehension measure, three for each time point. We computed scores as the number of correct answers. We divided the stories into three groups of similar difficulty to be presented at the three time points. We initially based the classification on a pilot sample of 13 monolingual children in Year 1. Because a second, larger sample (40 children in Year 1, of which 26 were bilingual) highlighted different difficulty levels between the local inference questions included in the stories at Time 2 and at Time 3, we eliminated outlier questions from the analyses, which resulted in our retaining seven local inference questions at Time 1 and Time 2 and five questions at Time 3 (see Appendix S4 in the online Supporting Information for details of piloting and question selection). Our listening comprehension measure showed good criterion validity (correlation with Understanding of Spoken Paragraphs subtest: Time 1, $r = .62$, $p < .001$; Time 3, $r = .75$, $p < .001$) and reliability (Time 1: Cronbach's $\alpha = .76$; Time 2: Cronbach's $\alpha = .82$; Time 3: Cronbach's $\alpha = .72$).

Parental Questionnaire

The children's parents completed a questionnaire (adapted from Serratrice & De Cat, 2020; see Appendix S6 in the online Supporting Information) on how much their children heard English and their home language, allowing us to extrapolate a measure of relative amount of English input, and provided demographic information relating to maternal education and socioeconomic status (SES). As a proxy for SES, we used the highest occupation in the household following the Standard Occupation Classification of the United Kingdom's

Office for National Statistics (https://onsdigital.github.io/dp-classification-tools/standard-occupational-classification/ONS_SOC_occupation_coding_tool.html). This classification yields lower scores for the higher earning occupations, thus, for clarity, we reversed the scores in the analyses.

To measure English input, we extrapolated a measure that considered the total amount of English input each child had received at Time 1 based on questionnaires that the parents completed between Time 1 and Time 2. Specifically, we computed the cumulative amount of English input as English exposure percentage multiplied by the length of exposure to English. We computed length of exposure to English as the number of months of exposure to English calculated from age of first exposure (i.e., the difference between age at Time 1 and age of first exposure). To compute English exposure in percentage, we asked the parents to indicate who spoke to the child, in which language, and how often over the course of the week using a 5-point scale, then we converted their responses into percentages: 0% (*never*), 25% (*rarely*), 50% (*half the time*), 75% (*usually*), 100% (*always*). The amount of input in English was the sum of the number of hours that a child spent with each interlocutor over the week multiplied by the percentage of time the child heard English from each interlocutor and then divided by the total number of hours the child was assumed to be awake, that is, 14 hours a day. For the time spent at school (6 hours a day), we assumed that the child heard only English. We recognize that this measure might only approximate the total exposure to English for each child because it assumed no change over the years, but we considered it to be a good enough approximation, given that we could not collect measures of relative amount of input at three time points due to resource limitations and the lack of direct access to the parents.

Data Analysis

We performed four analyses. First, we explored the relationship between measures using correlations, considering percentages of correct responses for listening comprehension.

Second, we used a repeated measures generalized linear model to explore predictors of listening comprehension at all three time points, considering each question separately for each participant, as the dependent measure. This simultaneously allowed us to consider different intercepts for each participant and for each question. We considered all the cognitive and language variables that correlated with the dependent variable as predictors in the model, as well as question type (i.e., comprehension of literal information, local and global inferencing) and time, and the interactions of question type, time, and all other significant predictors.

Third, to explore the relationship between time and question type in more depth, we modeled literal comprehension and local and global inferences separately.

Fourth, we explored the effect of cumulative amount of English input on listening comprehension. Specifically, we performed a mediation analysis to explore whether English input predicted listening comprehension (measured as percentage scores at each time point), either directly, or indirectly via other language skills (acting as mediators). We also considered maternal education, SES, and age at Time 1 as potential mediated predictors of listening comprehension because the literature has shown that these measures have an effect on language skills such as vocabulary and morphosyntax (Paradis, 2011; Unsworth, 2016), but none of these measures correlated significantly with listening comprehension at all time points, thus we did not perform mediation analyses for these measures. We performed analyses in R (R Core Team, 2021), and we used IBM SPSS for Windows (Version 25) for two principal component analyses, one of vocabulary depth measures, and one of memory measures. We applied a significance level of .05 in all analyses unless otherwise specified.

Results

Table 1 provides descriptive statistics for all the measures. To avoid issues of collinearity, we combined the measures of memory and of vocabulary depth at Time 1 to form two factors in a principal component analysis. For both factors, the correlations between the individual measures were significant and higher or equal to .30 (see Appendix S1 in the online Supporting Information).

No measure showed floor or ceiling effects except for nonverbal inferencing, which showed a significant positive skew because a quarter of the participants obtained a score of 9 out of 10. In the specially designed listening comprehension measure, the children showed progression over time; the children also showed higher scores in local inferences (Time 1: $M = 61\%$, $SD = 24$, 95% CI [56.3, 65.7]; Time 2: $M = 64\%$, $SD = 25$, 95% CI [59, 69]; Time 3: $M = 60\%$, $SD = 28$, 95% CI [54.2, 65.8]) than in literal questions (Time 1: $M = 44\%$, $SD = 24$, 95% CI [39.3, 48.7]; Time 2: $M = 55\%$, $SD = 24$, 95% CI [50.1, 59.9]; Time 3: $M = 68\%$, $SD = 23$, 95% CI [63.2, 72.8]) or in global inferences (Time 1: $M = 40\%$, $SD = 21$, 95% CI [35.9, 44.1]; Time 2: $M = 49\%$, $SD = 25$, 95% CI [44, 54]; Time 3: $M = 52\%$, $SD = 20$, 95% CI [47.8, 56.2]).

Table 1 Descriptive statistics and Spearman rho (nonnormally distributed variables) and Pearson (normally distributed variables) correlation coefficients between all measures (raw scores; $N = 100$ unless specified)

Variable	*M*	95% CI	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Maternal education ($N = 89$)	3.36	[3.19, 3.53]	–																
2. Socioeconomic status ($N = 74$)	6.43	[5.65, 7.21]	**.68**	–															
3. Age at Time 1 in months	68.40	[67.70, 69.10]	.08	.19	–														
4. Cumulative English input ($N = 92$)	25.18	[21.81, 28.55]	**.32**	**.42**	.20	–													
5. Nonverbal abilities	14.26	[13.62, 14.90]	.06	.26	.21	.25	–												

(*Continued*)

Table 1 (Continued)

Variable	*M*	95% CI	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
6. Attention: TROG: Score!	5.74	[5.31, 6.17]	.08	−.06	.14	.06	.06	–											
7. Attention: TROG: Walk Don't Walk	15.33	[14.55, 16.11]	.08	−.06	.05	.02	.05	**.35**	–										
8. Memory factor	0.00	[−0.12, 0.12]	.07	.01	.20	**.40**	.22	**.34**	**.28**	–									
Forward digit span	4.07	[3.90, 4.24]	–	–	–	–	–	–	–										
Backward digit span	2.28	[2.11, 2.45]	–	–	–	–	–	–	–										
Backward word total	4.00	[3.65, 4.35]	–	–	–	–	–	–	–										
Backward word span	1.91	[1.82, 2.00]	–	–	–	–	–	–	–										

(*Continued*)

Table 1 (Continued)

Variable	*M*	95% CI	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
9. Vocabulary breadth (BPVS)	68.83	[65.78, 71.88]	**.28**	**.38**	.13	**.64**	**.32**	**.35**	.15	**.52**[a]	–								
10. Vocabulary Depth factor	0.00	[−0.12, 0.12]	**.37**	**.33**	.21	**.47**	**.46**	**.34**	.23	**.53**[a]	**.66**[a]	–							
Synonyms	7.84	[6.80, 8.87]	–	–	–	–	–	–	–	–	–								
Opposites	9.17	[8.14, 1.20]	–	–	–	–	–	–	–	–	–								
Word Classes receptive	16.07	[15.46, 16.68]	–	–	–	–	–	–	–	–	–								
Word Classes expressive	1.58	[9.74, 11.42]	–	–	–	–	–	–	–	–	–								
11. Morphosyntax: Word Structures	14.53	[13.31, 15.75]	.18	.18	.15	**.48**	**.32**	.16	.19	**.55**	**.70**	**.62**	–						

(*Continued*)

Table 1 (Continued)

Variable	*M*	95% CI	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
12. Morphosyntax: TROG short	22.21	[2.64, 23.78]	**.29**	.23	.19	**.46**	**.42**	.22	.05	**.51**	**.69**	**.66**	**.66**	–					
13. Nonverbal inferencing	7.70	[7.36, 8.04]	.11	.11	.08	.07	.25	.05	.10	.19	**.36**	**.31**	**.41**	**.35**	–				
14. Comprehension monitoring	7.24	[6.91, 7.58]	.04	.06	.19	**.27**	.12	.12	.07	**.26**	**.26**	.25	.26	.24	.07	–			
15. List Comp Time 1 %	48.00	[44.50, 51.50]	.16	.22	.08	**.51**	**.30**	.01	.09	**.50**[a]	**.69**[a]	**.64**[a]	**.73**	**.56**	**.38**	.15	–		
16. List Comp Time 2 % ($N = 94$)	56.00	[51.90, 6.10]	.10	.12	.11	**.42**	**.27**	.03	.12	**.50**[a]	**.54**[a]	**.57**[a]	**.63**	**.56**	**.36**	.02	**.76**[a]	–	
17. List Comp Time 3 % ($N = 89$)	59.00	[55.50, 62.50]	**.30**	.22	.13	**.37**	**.32**	.05	.01	**.33**[a]	**.55**[a]	**.54**[a]	**.64**	**.62**	**.44**	.06	**.68**[a]	**.71**[a]	–

Note. All coefficients are Spearman rho unless otherwise noted. Correlations were not computed for the components of the two factors. Coefficients in boldface are significant at $\alpha = .01$. TROG = Test of Reception of Grammar; BPVS = British Picture Vocabulary Scale; List Comp = listening comprehension.
[a]Pearson correlation coefficient.

We did not use standard scores in the analyses, however, we computed standard scores at Time 1 for a general overview of the children's skills. As we had expected, the children's receptive vocabulary skills were lower than monolingual norms, with nine children with a standard score lower than 70 in vocabulary breadth (British Picture Vocabulary Scale, Dunn et al., 2009, $M_{\text{standard score}} = 87.17$; $SD = 11.62$). However, the children performed within the normal range in the test of nonverbal reasoning (Matrix Reasoning from the Wechsler Preschool and Primary Scale of Intelligence, Wechsler, 2013, $M_{\text{scaled score}} = 9.44$, $SD = 2.31$). The mean cumulative English input computed from the parental questionnaire was 25 months, corresponding to around two years equivalent of monolingual exposure; the children varied widely in this measure, with 25 children with a cumulative English input score lower than 12 months, 36 children with a cumulative English input score between 12 and 30 months (one to two years and a half), and 13 with a cumulative English input score higher than 48 months (four years).

Table 1 presents correlations between all measures. We applied a restrictive significance level of .01 to adjust for the large number of correlations performed. As we had expected, all the measures of foundational language skills correlated moderately to highly with each other. The Memory factor and the measure of nonverbal abilities also moderately correlated with all the lower level language measures. The two measures of attention, on the other hand, showed significant correlation only between themselves and memory skills, except a low correlation between Score! and the Vocabulary Depth factor. Scores on comprehension monitoring were not at chance, $W = 2{,}973.50$, $p < .001$, but the task showed no significant relationship with other measures except low correlations with English input, the Memory factor, and the vocabulary breadth measure. Nonverbal inferencing showed weak, but significant, correlations with all foundational language measures. Maternal education correlated with SES, English input and the Vocabulary Depth factor. SES correlated with English input and both measures of vocabulary. English input correlated with all foundational measures of language, while the age of the participants did not correlate with any other measure.

Table 1 also shows significant correlations between listening comprehension and nonverbal abilities, the Memory factor, vocabulary breadth, the Vocabulary Depth factor, both measures of morphosyntactic knowledge, and nonverbal inferencing. Thus we considered only these measures further in the models. We included English input in mediation models as a predictor of both lower level language skills and listening comprehension due to its correlation with these measures. We did not further include maternal education because it showed a correlation only with listening comprehension at Time 3.

Predictors of Listening Comprehension at All Time Points

We conducted generalized linear mixed-effects models for binomial data (Jaeger, 2008) using the software (R Core Team, 2021) and the function glmer from the package lme4 (Bates, Maechler, Bolker, & Walker, 2014). The children provided scores for six literal questions at each time point, scores for seven local inferences at Time 1 and Time 2 and for five local inferences at Time 3, and scores for nine global inference items at each time point. The scoring for each item was binary: 1 for a correct response and 0 for an incorrect response. Our dependent variable was the binary scoring for each item for each participant at each time point. All 100 participants provided scores at Time 1, 94 provided scores at Time 2, and 89 provided scores at Time 3. We used time as a continuous variable (Grimm, Ram, & Estabrook, 2017) and centered it around the second time point. As such, the intercept in the following models represents the listening comprehension score at Time 2.

Because we hypothesized growth in listening comprehension skills over time, an initial model included the random intercept terms for both participants and question to account for participant- and question–related variation at Time 2. We also included the random slope of participant over time to model participant-related variation in the effect of the repeated measures, together with the fixed effect of time. Comparing this model with a model that postulated no growth (i.e., a model that included only the random intercepts) confirmed linear growth for listening comprehension scores over time, $\chi^2(3) = 70.67$, $p < .001$. However, the comparison between a model with and without random slope highlighted no random slope effect, $\chi^2(2) = 0.48, p = .785$, suggesting that, although the children showed growth in their listening comprehension scores, as shown by the significant fixed effect of time, this growth was similar for all participants. All further analyses retained a random intercept term for both participants and question but no random slope of participant over time: we use this model that included only the fixed effect of time as a baseline model (henceforth, time-only model).

We then compared this time-only model (using pairwise likelihood ratio test comparisons; Barr, Levy, Scheepers, & Tily, 2013) with a model that additionally included the hypothesized fixed effects: question type (literal vs. local inference vs. global inference; literal questions as reference level), nonverbal abilities, the Memory factor, vocabulary breadth, the Vocabulary Depth factor, the measures of morphosyntactic knowledge and nonverbal inferences. We centered all the continuous variables around the sample mean, except those for the two factors created through principal component analysis (Memory and Vocabulary Depth). This model improved fit compared to the time-only model, $\chi^2(9) = 249.17, p < .001$.

We included the hypothesized interactions of time and any significant fixed effect one at a time in the model with all fixed effects, but no interactions were significant. We separately introduced the interactions of type of question and the measures of cognitive and language abilities that were significant to test whether individual differences in cognitive and language abilities differentially predicted different types of questions. We reduced nonconvergent models to significant predictors only to facilitate convergence. Neither the Vocabulary Depth factor, $\chi^2(2) = 5.21, p = .074$, nor word structure, χ^2 (2) = 4.13, $p =$ .127, showed a significant interaction with question type.

To further explore whether the three types of questions developed similarly over time, we added the interaction of type of question with time to the model. The addition of the interaction improved model fit, $\chi^2(2) = 60.25$, p $< .001$. Table 2 presents the results of the final model. Figure 1 represents the interaction between time and question type. The final model showed no issue of collinearity; the variance inflation factor for time and the interaction Time × Question Type was smaller than 5, and all other variance inflation factors smaller than 3.

As Table 2 shows, the Vocabulary Depth factor and morphological knowledge measured through the Word Structures subtest significantly predicted listening comprehension. Question type was also a significant predictor, with local inferences emerging as the easiest questions, above literal questions and global inference questions, which emerged as the most difficult type of questions. The results also highlight different trajectories for the different questions (Figure 1) with higher growth over time in comprehension for literal questions, followed by growth in comprehension for global inferences questions, and a leveling off over time for the comprehension of local inferences questions. The model explained 30% of variance in listening comprehension—this level of explained variance is to be expected for logistic models; of this 30%, the fixed effects explained 17%. Although fairly low compared to values normally expected for simpler regression models (Moore, Notz, & Flinger, 2013), this value was still much higher than the variance explained by time in the time-only model (1%).

Table 2 Results for the repeated measures generalized linear mixed effects model

Parameter	*b*	*SE*	95 % CI	*OR*	*z*	*p*	χ^2	*df*	*p*
Intercept	0.42	0.16	[0.11, 0.72]	1.52	2.69	**.007**			
Time	0.76	0.08	[0.61, 0.92]	2.14	9.64	**< .001**	70.72	1	**< .001**
Question type							130.60	2	**< .001**
Literal vs. Local	−0.08	0.02	[−0.12, −0.04]	0.92	−3.41	**.002**			
Literal vs. Global	0.15	0.03	[0.09, 0.21]	1.16	6.03	**< .001**			
Local vs. Global	0.23	0.02	[0.19, 0.27]	1.26	11.98	**< .001**			
Nonverbal abilities	−0.03	0.06	[−0.14, 0.09]	0.97	−0.46	.642	0.28	1	.597
Vocabulary									
Breadth(BPVS)	0.09	0.08	[−0.08, 0.26]	1.09	1.07	.285	1.15	1	.283
Depth factor	0.22	0.08	[0.07, 0.38]	1.25	2.85	**.004**	7.41	1	**.006**
Morphosyntax									
Word Structures	0.44	0.08	[0.28, 0.60]	1.55	5.34	**< .001**	22.29	1	**< .001**
TROG short	0.01	0.08	[−0.14, 0.17]	1.01	0.16	.870	0.04	1	.846
Nonverbal inferences	0.06	0.06	[−0.06, 0.38]	1.06	0.98	.329	1.02	1	.313
Question Type × Time							61.22	2	**< .001**
Literal vs. Local	0.96	0.13	[0.70, 1.22]	2.62	7.70	**< .001**			
Literal vs. Global	0.37	0.10	[0.17, 0.57]	1.44	3.53	**.012**			
Local vs. Global	−0.60	0.12	[−0.84, −0.36]	0.55	−4.98	**< .001**			
Random effects	**Variance**	***SD***		$R^2_{marginal}$	$R^2_{conditional}$				
Subject	0.19	0.43		.17	.30				
Question	0.40	0.63							

Note. Given convergence issues, the final model did not include the Memory factor because the effect of this variable was not significant and had the lowest estimated effect size. Values in boldface are significant for $\alpha = .05$. BPVS = British Picture Vocabulary Scale; TROG = Test of Reception of Grammar.

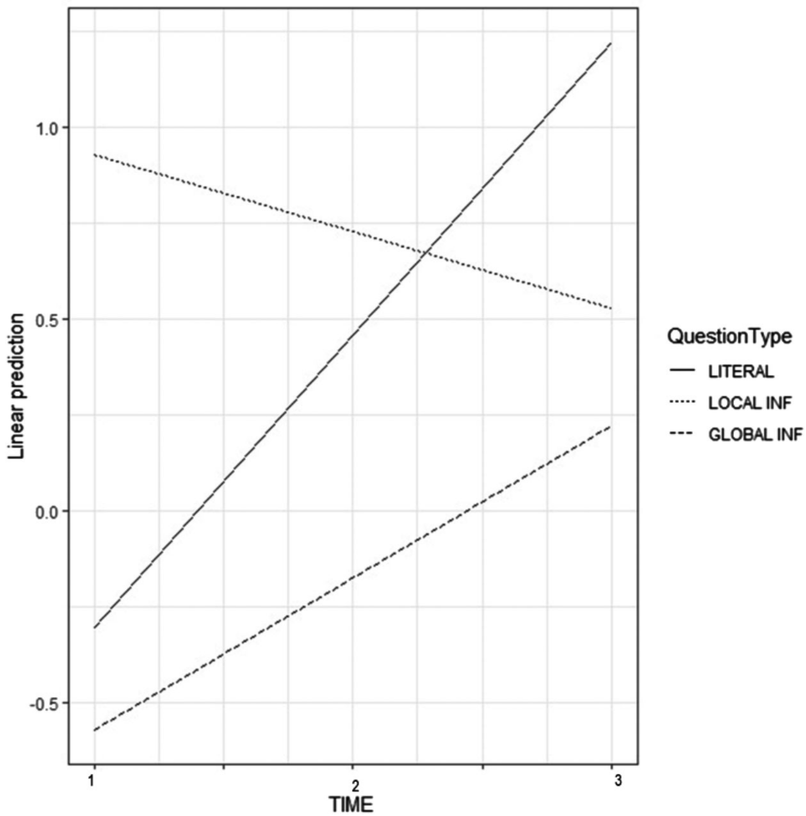

Figure 1 Linear prediction of listening comprehension by time and question type. INF = inference.

Predictors and Differences in Growth for Each Question Type

Given the significant interaction of time with question type, we computed three models, one for each type of question. We used these models to explore growth in comprehension for each type of question. We used the same predictors as the ones that we had entered in the main model to ensure comparability with the main model. We deemed these analyses appropriate given the significant interaction of question type with time.

The models (see Tables 3, 4, and 5) showed that comprehension of literal questions and global inference questions significantly improved over time, but comprehension of local inference questions did not. As in the main model, both the Vocabulary Depth factor and the morphological knowledge measure

Table 3 Results of the repeated measures generalized linear mixed effect models for literal questions

Parameter	*b*	95 % CI	*OR*	z	*p*
Intercept	0.67	[−0.13, 1.47]	1.95	1.64	.101
Time	1.12	[0.90, 1.34]	3.07	10.04	**< .001**
Nonverbal abilities	−0.04	[−0.22, 0.14]	0.96	−0.45	.652
Memory factor	0.04	[−0.17, 0.24]	1.04	00.35	.727
Vocabulary					
Breadth (BPVS)	−0.11	[−0.36, 0.14]	0.89	−0.89	.375
Depth factor	0.14	[−0.10, 0.37]	1.15	1.14	.256
Morphosyntax					
Word Structures	0.67	[0.42, 0.93]	1.96	5.07	**< .001**
TROG short	0.10	[−0.14, 0.34]	1.11	0.82	.415
Nonverbal inferences	0.03	[0.14, 0.20]	1.03	0.35	.730
Random effects	**Variance**	***SD***	$R^2_{marginal}$	$R^2_{conditional}$	
Subject	0.18	0.42	.20	.52	
Question	0.40	0.63			

Note. Values in boldface are significant for $\alpha = .05$. BPVS = British Picture Vocabulary Scale; TROG = Test of Reception of Grammar.

predicted comprehension of both types of inference questions, but only morphological knowledge predicted literal questions. Figure 2 shows the children's comprehension of literal questions and of local and global inference questions when we grouped the children by morphological knowledge or the Vocabulary Depth factor. Higher ability children showed higher scores but relatively similar growth (or lack of growth) over time compared to low ability children in all question types, except for an apparent lack of growth in global inference questions between Time 2 and Time 3 for the top quartile group in morphological knowledge and the two higher groups for vocabulary depth.

Mediation Model

English input was the only environmental measure (among English input, maternal education, and SES) to significantly correlate with listening comprehension at all time points. English input also correlated with all other language measures (see Table 1). To explore the effect of English input on

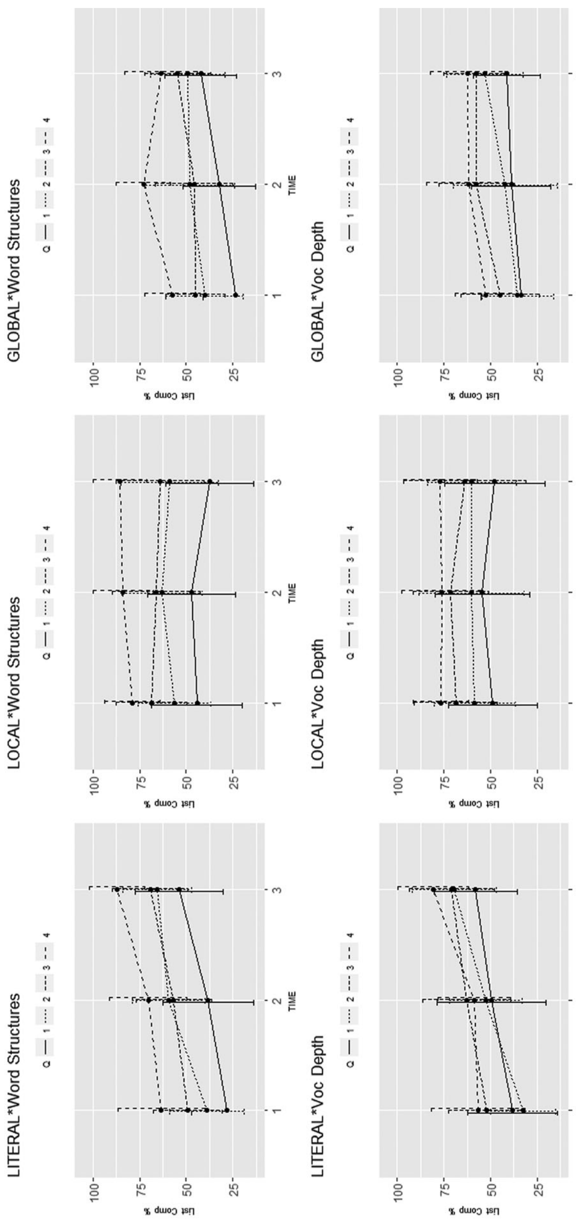

Figure 2 Listening comprehension % scores distribution (with *SD*) for each question type (literal, local, and global inferences from left to right) at each time point (within graphs). Top row: participants divided by score for Word Structures subtest in quartiles (Q). Bottom row: participants divided by score for Vocabulary Depth factor (Voc Depth) in quartiles (Q).

Table 4 Results of the repeated measures generalized linear mixed effect models for local inferences

Parameter	*b*	95 % CI	*OR*	z	*p*
Intercept	0.58	[0.28, 0.88]	1.79	3.84	**< .001**
Time	−0.03	[−0.29, 0.24]	0.97	−0.20	.840
Nonverbal abilities	−0.01	[−0.15, 0.15]	1.00	−0.02	.982
Memory factor	0.01	[−0.16, 0.19]	1.01	0.16	.872
Vocabulary					
Breadth (BPVS)	0.13	[−0.09, 0.35]	1.14	1.13	.260
Depth factor	0.27	[0.07, 0.48]	1.31	2.58	**.009**
Morphosyntax					
Word Structures	0.51	[0.28, 0.73]	1.66	4.44	**< .001**
TROG short	−0.06	[−0.26, 0.15]	0.95	−0.53	.599
Nonverbal inferences	0.08	[−0.07, 0.23]	1.08	1.04	.297
Random effects	**Variance**	***SD***	$R^2_{marginal}$	$R^2_{conditional}$	
Subject	0.18	0.42	.15	.25	
Question	0.28	0.53			

Note. Values in boldface are significant for $\alpha = .05$. BPVS = British Picture Vocabulary Scale; TROG = Test of Reception of Grammar.

listening comprehension while accounting for its possible effect on other language measures, we computed a mediation model. We specifically considered as potential mediators those language variables that were significant in the main mixed-effects model (word structures and the Vocabulary Depth factor). The mediation model was built using the R package lavaan (Rosseel, 2012). As the dependent variable in the model, we used a latent variable directly estimated within the model, computed considering percentages of correct responses in listening comprehension at each time point. We estimated regression paths between listening comprehension at Time 1 and Time 2 and between Time 2 and Time 3 to further consider the growth of listening comprehension between time points shown in the main model. We computed fit for the model using maximum likelihood estimation with robust standard errors.

The model (see Figure 3) explained 78.2% of the variance in listening comprehension. It highlighted no significant direct effect of English input on listening comprehension but significant indirect effects of English input through

Table 5 Results of the repeated measures generalized linear mixed effect models for global inferences

Parameter	*b*	95 % CI	*OR*	z	*p*
Intercept	−0.13	[−0.55, 0.29]	0.88	−0.61	.539
Time	0.40	[0.25, 0.56]	1.49	5.05	**< .001**
Nonverbal abilities	−0.04	[−0.18, 0.11]	0.96	−0.51	.613
Memory factor	−0.03	[−0.19, 0.14]	0.97	−0.33	.744
Vocabulary					
Breadth (BPVS)	0.20	[−0.01, 0.40]	1.22	1.90	.057
Depth factor	0.27	[0.08, 0.46]	1.31	2.73	**.006**
Morphosyntax					
Word Structures	0.32	[0.11, 0.53]	1.37	3.01	**.003**
TROG short	0.01	[−0.19, 0.21]	1.01	0.10	.921
Nonverbal inferences	0.07	[−0.07, 0.21]	1.07	1.00	.319
Random effects	**Variance**	***SD***	$R^2_{marginal}$	$R^2_{conditional}$	
Subject	0.19	0.43	.12	.32	
Question	0.40	0.63			

Note. Values in boldface are significant for $\alpha = .05$. BPVS = British Picture Vocabulary Scale; TROG = Test of Reception of Grammar.

the Word Structures subtest and the Vocabulary Depth factor, as well as significant effects of the mediators, confirming the results of the model in Table 2: Word Structures, $b = 0.63$, $p < .001$;Vocabulary Depth, $b = 0.29$, $p = .005$. The total effect of English input on listening comprehension was Path c_1, $b = 0.50$, $p < .001$.

Discussion

The aim of our study was to investigate how foundational cognitive and language skills predict literal and inferential listening comprehension over time in bilingual children between the ages of 5 and 7 years (Year 1 to Year 2 of primary school in the English school system) and how this relationship is mediated by English input. We also addressed whether the comprehension of literal information and of local and global inferences is differentially affected by foundational cognitive and language skills in bilingual children.

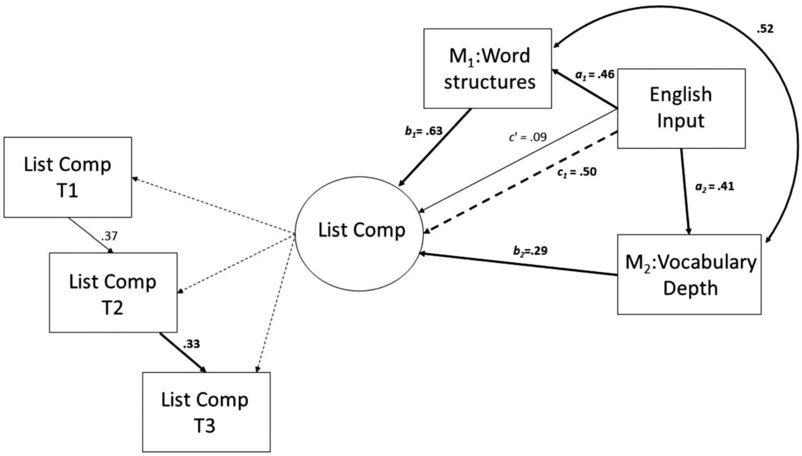

Figure 3 Multiple mediation model with two mediators: M_1, Word Structures, and M_2, Vocabulary Depth factor, $\chi^2(4) = 0.93$, $p = .920$, comparative fit index > .999, Tucker Lewis index > .999, root mean square error of approximation < .001, standardized root mean square residual = .011. English input had no significant direct effect: Path c', $b^* = 0.09$, 95% CI [−0.07, 0.23], $p = .263$. English input had a significant indirect effect on listening comprehension through M_1, computed as the product of the two paths linking English input on listening comprehension through that mediator, that is, Path a_1b_1, $b^* = 0.29$, 95% CI [0.15, 0.44], $p = .001$. Similarly, English input had a significant indirect effect on listening comprehension through M_2 (Vocabulary Depth factor), defined as Path a_2b_2, $b^* = 0.12$, 95% CI [0.03, 0.21], $p = .017$. The total effect of English input on listening comprehension was computed as: Path $c_1 = c' + a_1b_1 + a_2b_2$, $b^* = 0.50$, 95% CI [0.33, 0.66], $p < .001$.

Predictors of Listening Comprehension

Our results clearly showed the importance of vocabulary depth and morphological knowledge in listening comprehension: Of all the predictors, only these two significantly contributed to explaining the variance in listening comprehension abilities. The importance of measures of vocabulary and grammar is in line with previous research (Kim, 2016; Silva & Cain, 2015). These two variables had a direct effect on listening comprehension, but their interaction with time was not significant, thus vocabulary depth and morphological knowledge predicted listening comprehension abilities overall (i.e., the intercept of a growth model) but not change over time; that is, all children improved in comprehension over time similarly, regardless of their initial morphological and vocabulary depth skills. The Vocabulary Depth factor and morphological skills also predicted all three types of comprehension questions similarly (literal, local, and global), although Vocabulary Depth failed to reach significance in the model for literal comprehension. Different from some of the previous research (Kim, 2016), our study did not highlight a direct effect of memory or nonverbal inferencing skills on listening comprehension once we had considered all other predictors.

In terms of vocabulary knowledge, our findings suggest that the quality of lexical representations (vocabulary depth) is more informative for listening comprehension than is the sheer number of words children know (vocabulary breadth). This result is in line with the result of Strasser and Del Rio (2014), who showed that the effect of vocabulary breadth tends to be fully mediated by vocabulary depth in the concurrent prediction of listening comprehension of preschool children. The importance of vocabulary depth over vocabulary breadth is also in line with the lexical quality hypothesis (Perfetti & Hart, 2002) because vocabulary depth tasks tend to tap into lexical quality, whereas vocabulary breadth tasks do not. In Perfetti and Hart's view, the link between different aspects of word knowledge is extremely important. They defined a representation of high quality as one that includes information regarding different features of the same word, that is, the representation of form and meaning, as well as morphological information and information regarding its use, and they suggested that a high-quality representation of a word allows children to rapidly access all features of a given word. High-quality lexical representations facilitate word processing, thus freeing processing resources, and they also provide detailed and extensive semantic information that is necessary for local and global inferences.

The other significant predictor of listening comprehension in all of our analyses was morphological knowledge. Morphological knowledge emerged as a better predictor than syntactic knowledge in our study. As in previous research (Florit et al., 2013), we found that a sentence-picture matching task was not associated with text comprehension. The measure of morphological knowledge adopted in this study, however, was heavily linked to word knowledge, and it captured the children's ability to manipulate inflectional and derivational endings and to use appropriate anaphoric forms. The importance of morphological knowledge for local inferences, especially those that require linking pronouns with their antecedents, is relatively straightforward. Children rely on their knowledge of pronouns to solve anaphoric local inferences; however the results suggest a more widespread effect of morphological knowledge on all types of listening comprehension questions. The importance of morphological knowledge is in line with previous research (Babayiğit, 2014), although in many previous studies the effect of morphology has often been conflated with that of syntactic knowledge (e.g. Babayiğit, 2014; Muter et al., 2004) and has not been consistently replicated (Alonzo et al., 2016). Well-developed morphological knowledge contributes to higher lexical quality (Perfetti & Hart, 2002), and it is a powerful tool that children can rely on to make those text connections leading to successful inferences.

Our results also showed that, although vocabulary depth and morphological knowledge both had an effect on listening comprehension, they did not seem to affect its growth over time. This result is similar to that of Lepola et al. (2012), who found an effect of vocabulary on listening comprehension concurrently in monolingual Finnish 4-year-olds, but not longitudinally two years later. In their path analysis, which included the autoregressor of listening comprehension (i.e., how much listening comprehension at earlier time points explained listening comprehension at subsequent time points), vocabulary explained the level of listening comprehension at Time 1 but did not further explain its growth from Time 1 to Time 3 above what had originally been explained by the autoregressor. Similarly, Proctor, Silverman, Harring, and Montecillo (2012) showed an effect of both vocabulary breadth and morphological knowledge in predicting initial levels of reading comprehension but not the change in reading comprehension of monolingual and bilingual children in primary school over six months. Although vocabulary and morphological knowledge are important in determining children's performance in listening comprehension tasks, they do not necessarily predict developmental progress. Because we did not consider growth in vocabulary and morphological skills over time in the present article, we cannot conclude whether growth in either of these skills might predict growth in listening comprehension.

The Vocabulary Depth factor was significant in the main model, yet further analyses failed to show an effect of vocabulary depth on literal comprehension. A reduced role of vocabulary on literal comprehension compared to inferencing is in line with previous research (Cain & Oakhill, 2014). A large and deep lexicon may be more important for making connections within the text rather than for verbatim recall as is the case in literal comprehension. Breadth and depth of vocabulary can also assist in making global inferences that require extensive background knowledge. For example, given a scene where people swim in the water and build sandcastles, children who have a deeper and more connected semantic network will have a link between the words sandcastle, water, and beach, which will facilitate the inference that the scene takes place on a beach.

A striking result of our research was the lack of individual growth slopes in listening comprehension over time. Research on growth in comprehension skills does not always include random slopes (Raudszus, Segers, & Verhoeven, 2021) because individual variance in growth as captured by random slopes might account for the same variance as individual difference measures researchers wish to model. Our result could be interpreted by considering that we tested children who had only started attending primary school. Starting school will reduce the individual variation in the amount of input bilingual children receive in each language because they will all be exposed to the same amount of English during school hours. English input at school would therefore have acted as a leveler for our bilingual children whose exposure to the language prior to the start of formal education had varied widely. Additional support for this hypothesis comes from the absence of individual differences in the growth patterns in vocabulary breadth in the same sample of children (Valentini & Serratrice, 2021).

Another important result of our study is the lack of a significant association between all other skills measured and listening comprehension, once we had accounted for the Vocabulary Depth factor and morphological knowledge. Two of these measures—memory skills and inferential skills—have been previously shown to be directly associated with listening comprehension (Kim, 2016). The direct effect of memory has not been replicated in all studies (Lepola et al., 2012; Silva & Cain, 2015) and could be explained by considering that memory might have an indirect association with listening comprehension skills through its association with other foundational skills (Cain et al., 2004; de Bree & Zee, 2020). The difference between our study and previous ones (Florit et al., 2011; Kim, 2016) regarding the lack of influence of inferencing skills on listening comprehension might be ascribed to the specific measures used. Previous research that found a more prominent effect of inferencing skills on listening comprehension used inferencing tasks that relied heavily on children's language abilities. The similarities between the two tasks might therefore have increased the likelihood of finding a direct relationship between inferencing and listening comprehension. In our study, however, we specifically chose a measure of nonverbal inferencing that allowed us to measure our bilingual children's inferencing skills without an excessive reliance on their oral language abilities. This allowed us to explore the relationship between inferencing skills and listening comprehension more directly. It is also possible that, because we measured foundational language skills more thoroughly (vocabulary breadth and depth and morphological and syntactic abilities), the language tasks that measured foundational language skills in our study accounted for part of the variability explained by inferencing skills tasks in other studies. This would further confirm the importance of including measures of vocabulary depth and morphological knowledge when researchers explore predictors of listening comprehension.

Differences Between Types of Comprehension Questions

The children's literal comprehension and their ability to answer global inference questions grew over time, but their ability to answer local inference questions remained relatively stable. Furthermore, local inferences emerged as the easiest type of questions for the children, followed by literal questions and by global inference questions. This result seems at odds with those of previous literature that reported literal questions as the easiest types of questions (Alptekin & Erçetin, 2011; Cain & Oakhill, 2014). This difference might be due to the type of questions that we used and the level of detail that our study required: We provided all the information necessary to answer literal question directly in the text, but the children needed to correctly encode and retrieve this information to answer these questions. It is possible that our test of comprehension required a higher level of detail than other studies have required, thus making literal questions harder. However, the difference between local and global inferences, and the emergence of global inference questions as the more complex type of comprehension questions, is in line with previous research with monolingual children when they were reading (Cain & Oakhill, 2014), but not when they were listening (Currie & Cain, 2015; Freed & Cain, 2017). Global inference questions require the ability to use previous knowledge to extract a deeper meaning from a given text than a literal interpretation might provide. This process is demanding, and this task might be particularly hard for bilingual children in their second language (Hara & Tappe, 2016; Schönpflug & Küpping-Faturikova, 2020). Our bilingual children needed to retrieve information that might not have been stored in their second language or that might have been encoded differently in the two languages (Adams, 2016).

As for local inferences, although the children's performance was not at ceiling at Time 1, it is possible that the results at Time 1 were already so high that the children were left with relatively little room for improvement. It is also possible that the skills required to answer local inference questions develop before the time window examined and that no major changes take place between 5 and 7 years of age. Almost half of the questions categorized as local inferences (seven out of 19) required the anaphoric resolution of a third person singular subject pronoun (*he, she*). English-speaking monolingual children as young as 3 years of age can use gender information to find the correct antecedent of a subject pronoun (Arnold, Brown-Schmidt, & Trueswell, 2007), but little is known about bilingual children's anaphora resolution in English (see Serratrice & Hervé, 2015, for an overview). Success on local inferences in our study required the children to find an antecedent for an anaphora, either a noun or a pronoun, in the preceding sentence. Whenever the children had to find an antecedent for a pronoun, knowledge of grammatical gender would unambiguously lead them to the correct choice. Similarly, when finding an antecedent for a nominal anaphora, there was only one semantically plausible antecedent. Anaphora resolution can be a very complex task when contextual and semantic information increase the ambiguity of a potential antecedent, but it can also be a rather mechanistic process when there is little ambiguity, as in the stories used in our study. Conversely, the ability to make global inferences requires children to recruit information from long-term memory that grows as a function of their increasing experience of the world, and therefore growth over time is to be expected.

The Effect of English Input

A key result of our study is the effect of the cumulative amount of English input on listening comprehension. Specifically, English input showed a significant but indirect effect on listening comprehension mediated by morphological knowledge and vocabulary depth. This result confirms the importance of language input for foundational language skills such as vocabulary and morphosyntactic abilities (Brinchmann, Braeken, & Lyster, 2019; Hoff et al., 2018). The cumulative amount of English input for the bilingual children in this study did not directly affect listening comprehension, but it directly predicted levels of vocabulary and grammar skills, that, in turn, positively influenced the children's ability to comprehend texts.

Limitations

A limitation of our research is that, contrary to some models in previous research (Kim, 2016), we did not consider the possible mediation effect of higher language abilities (e.g., inferencing and comprehension monitoring) on the relationship between listening comprehension and cognitive skills. In addition, although we did include lower level cognitive skills (e.g. memory and attention), our results show the importance of vocabulary depth and morphological knowledge in predicting listening comprehension, over and above the influence of any other predictor. However, it is possible that lower level cognitive skills, such as memory and attention measured in our study, might have a subtler relationship with listening comprehension, mediated via a relationship between these skills and vocabulary and grammar skills. It could be, for instance, that children with better working memory might be better word learners, with better vocabulary skills, and that better vocabulary skills will positively affect their listening comprehension. Our initial choice of a simpler model that did not consider this direction of influence was motivated by the relatively lower number of participants in our study compared to studies that have considered these relationships (de Bree & Zee, 2020; Kim, 2016). We believe our model is of value in highlighting the specific importance of vocabulary and morphological skills in predicting listening comprehension, however we cannot rule out the possible (mediated) effects of other cognitive skills.

Another limitation of our research is the lack of longitudinal measurement of the predictors of listening comprehension as well as English input; specifically, our model considered how the children's abilities at the beginning of the study related to their listening comprehension longitudinally. We chose this approach due to lack of longitudinal data on some of the predictors. It is possible that, although initial skills might only predict level of listening comprehension, considering growth in these skills over time might have also predicted growth in listening comprehension. Our study is still novel in attempting to measure many of the possible predictors of growth in listening comprehension because many of the previous studies only considered listening comprehension concurrently (de Bree & Zee, 2020; Kim, 2016) or restricted their analysis to some predictors (Lepola et al., 2012), or both (Florit et al., 2011). However, we acknowledge that including longitudinal measures of the predictors, as well as longitudinal changes in English input, might have accounted for more variability in listening comprehension, especially in relation to its growth over time.

Conclusion

Vocabulary depth and morphological knowledge were the most significant predictors of English listening comprehension in the bilingual children in the first two years of formal schooling in our study. These skills specifically determined the children's listening comprehension but not their growth in listening comprehension abilities over time. The amount of English input that the children had received had a significant impact on their listening comprehension performance, which was mediated by its effect on foundational language skills, particularly morphological knowledge and the Vocabulary Depth factor. These results make a novel contribution to a better understanding of the determinants of listening comprehension in bilingual children, and they have pedagogical implications. Increasing children's high-quality lexical representations, specifically providing them with more information about the meaning and use of words as well as increasing their knowledge of morphology, is likely to have a positive, cascading effect on their spoken language comprehension.

Regarding different kinds of comprehension questions, the results confirm that those whose answers require global inferences are the hardest kinds of questions for school-aged children and that both their ability to answer global inference questions and their literal comprehension grow over time.

Final revised version accepted 11 March 2022

Open Research Badges

This article has earned Open Data badges for making publicly available the digitally-shareable data and the components of the research methods needed to reproduce the reported procedure and results. All data that the authors have used and have the right to share are available at https://osf.io/2fa3c/?view_only=cc167f7a3484432b9d5d56e4b87d4032 and IRIS http://www.iris-database.org.

References

Adams, A. M. (2016). How language is embodied in bilinguals and children with specific language impairment. *Frontiers in Psychology*, *7*, 1209. https://doi.org/10.3389/fpsyg.2016.01209

Alonzo, C. N., Yeomans-Maldonado, G., Murphy, K. A., Bevens, B., & Language and Reading Research Consortium. (2016). Predicting second grade listening comprehension using prekindergarten measures. *Topics in Language Disorders*, *36*, 312–333. https://doi.org/10.1097/tld.0000000000000102

Alptekin, C., & Erçetin, G. (2011). Effects of working memory capacity and content familiarity on literal and inferential comprehension in L2 reading. *TESOL Quarterly*, *45*, 235–266. https://doi.org/10.5054/tq.2011.247705

Arnold, J. E., Brown-Schmidt, S., & Trueswell, J. (2007). Children's use of gender and order-of-mention during pronoun comprehension. *Language and Cognitive Processes*, *22*, 527–565. https://doi.org/10.1080/01690960600845950

Babayiğit, S. (2014). The role of oral language skills in reading and listening comprehension of text: A comparison of monolingual (L1) and bilingual (L2) speakers of English language. *Journal of Research in Reading*, *37*(S1), S22–S47. https://doi.org/10.1111/j.1467-9817.2012.01538.x

Babayiğit, S., & Shapiro, L. (2020). Component skills that underpin listening comprehension and reading comprehension in learners with English as first and additional language. *Journal of Research in Reading*, *43*, 78–97. https://doi.org/10.1111/1467-9817.12291

Baddeley, A. D. (1986). *Working memory*. Oxford, UK: Oxford University Press

Barnes, M. A., Dennis, M., & Haefele-Kalvaitis, J. (1996). The effects of knowledge availability and knowledge accessibility on coherence and elaborative inferencing in children from six to fifteen years of age. *Journal of Experimental Child Psychology*, *61*, 216–241. https://doi.org/10.1006/jecp.1996.0015

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates D., Maechler M., Bolker B., &Walker S. (2014). lme4: Linear mixed-effects models using Eigen and S4 (R package, Version 1.1-7) [Computer software]. Retrieved from https://cran.r-project.org/src/contrib/Archive/lme4/lme4_1.1-7.tar.gz

Bishop, D. V. M. (2003). Test for the Reception of Grammar *(Version 2) [Assessment instrument*]. London, UK: Harcourt Assessment.

Blom, E. (2010). Effects of input on the early grammatical development of bilingual children. *International Journal of Bilingualism*, *14*, 422–446. https://doi.org/10.1177/1367006910370917

Bowyer-Crane, C., Fricke, S., Schaefer, B., Lervåg, A., & Hulme, C. (2017). Early literacy and comprehension skills in children learning English as an additional language and monolingual children with language weaknesses. *Reading and Writing*, *30*, 771–790. https://doi.org/10.1007/s11145-016-9699-8

Brinchmann, E. I., Braeken, J., & Lyster, S. A. H. (2019). Is there a direct relation between the development of vocabulary and grammar? *Developmental Science*, *22*(1), e12709. https://doi.org/10.1111/desc.12709

Burgoyne, K., Kelly, J. M., Whiteley, H. E., & Spooner, A. (2009). The comprehension skills of children learning English as an additional language. *British Journal of Educational Psychology*, *79*, 735–747. https://doi.org/10.1348/000709909x422530

Cain, K., & Oakhill, J. (2014). Reading comprehension and vocabulary: Is vocabulary more important for some aspects of comprehension? *L'Annee Psychologique*, *114*, 647–662. https://doi.org/10.4074/s0003503314004035

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*, 31–42. https://doi.org/10.1037/0022-0663.96.1.31

Currie, N. K., & Cain, K. (2015). Children's inference generation: The role of vocabulary and working memory. *Journal of Experimental Child Psychology*, *137*, 57–75. https://doi.org/10.1016/j.jecp.2015.03.005

de Bree, E., & Zee, M. (2020). The unique role of verbal memory, vocabulary, concentration and self-efficacy in children's listening comprehension in upper elementary grades. *First Language*, *41*, 129–153. https://doi.org/10.1177/0142723720941680

Department for Education (2020). *English proficiency of pupils with English as an additional language. Ad-hoc notice*. Retrieved from:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/868209/English_proficiency_of_EAL_pupils.pdf
Dixon, C., Thomson, J., & Fricke, S. (2020). Language and reading development in children learning English as an additional language in primary school in England. *Journal of Research in Reading*, *43*, 309–328. https://doi.org/10.1111/1467-9817.12305
Dunn, L. M., Dunn, D. M., & NFER. (2009). *British Picture Vocabulary Scale* (3rd ed.) [Assessment instrument]. London, UK: GL Assessment.
Eason, S. H., Goldberg, L. F., Young, K. M., Geist, M. C., & Cutting, L. E. (2012). Reader–text interactions: How differential text and question types influence cognitive skills needed for reading comprehension. *Journal of Educational Psychology*, *104*, 515–528. https://doi.org/10.1037/a0027182
Florit, E., Roch, M., Altoé, G., & Levorato, M. C. (2009). Listening comprehension in preschoolers: The role of memory. *British Journal of Developmental Psychology*, *27*, 935–951. https://doi.org/10.1348/026151008x397189
Florit, E., Roch, M., & Levorato, M. C. (2011). Listening text comprehension of explicit and implicit information in preschoolers: The role of verbal and inferential skills. *Discourse Processes*, *48*, 119–138. https://doi.org/10.1080/0163853x.2010.494244
Florit, E., Roch, M., & Levorato, M. C. (2013). The relationship between listening comprehension of text and sentences in preschoolers: Specific or mediated by lower and higher level components. *Applied Psycholinguistics*, *34*, 395–415. https://doi.org/10.1017/s0142716411000749
Freed, J., & Cain, K. (2017). Assessing school-aged children's inference-making: The effect of story test format in listening comprehension. *International Journal of Language & Communication Disorders*, *52*, 95–105. https://doi.org/10.1111/1460-6984.12260
Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Valimaa, T., Balciuniene, I., … Walters, J. (2012). MAIN: Multilingual Assessment Instrument for Narratives. *ZAS Papers in Linguistics*, *56*. Berlin, Germany: Zentrum fur Allgemeine Sprachwissenschaft. https://doi.org/10.21248/zaspil.56.2019.414
Gibson, T. A., Oller, D. K., Jarmulowicz, L., & Ethington, C. A. (2012). The receptive-expressive gap in the vocabulary of young second-language learners: Robustness and possible mechanisms. *Bilingualism: Language and Cognition*, *15*, 102–116. https://doi.org/10.1017/s1366728910000490
Grimm, K. J., Ram, N., & Estabrook, R. (2017). *Growth modeling: Structural equation and multilevel approaches*. London, UK: Guiford.
Grüter, T., & Paradis, J. (Eds.). (2014). *Input and experience in bilingual development, Trends in language acquisition research* (Vol. *13*). Amsterdam, The Netherlands: John Benjamins.
Hammer, C. S., Komaroff, E., Rodriguez, B. L., Lopez, L. M., Scarpino, S. E., & Goldstein, B. (2012). Predicting Spanish–English bilingual children's language

abilities. *Journal of Speech, Language, and Hearing Research*, *55*, 1251–1264. https://doi.org/10.1044/1092-4388(2012/11-0016)

Hara, A., & Tappe, H. (2016). Inference generation and text comprehension in bilingual children: A case study. *Literator*, *37*(2), 1–12. https://doi.org/10.4102/lit.v37i2.1287

Hoff, E., Quinn, J. M., & Giguere, D. (2018). What explains the correlation between growth in vocabulary and grammar? New evidence from latent change score analyses of simultaneous bilingual development. *Developmental Science*, *21*(2), e12536. https://doi.org/10.1111/desc.12536

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, *2*, 127–160. https://doi.org/10.1007/bf00401799

Hutchinson, J. M., Whiteley, H. E., Smith, C. D., & Connors, L. (2003). The developmental progression of comprehension-related skills in children learning EAL. *Journal of Research in Reading*, *26*, 19–32. https://doi.org/10.1111/1467-9817.261003

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*, 236–248. https://doi.org/10.1037/0012-1649.27.2.236

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, *45*, 337–374. https://doi.org/10.1016/S0010-0285(02)00500-5

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434–446. https://doi.org/10.1016/j.jml.2007.11.007

Kim, Y. S. G. (2016). Direct and mediated effects of language and cognitive skills on comprehension of oral narrative texts (listening comprehension) for children. *Journal of Experimental Child Psychology*, *141*, 101–120. https://doi.org/10.1016/j.jecp.2015.08.003

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, *49*, 294–303. https://doi.org/10.1037/0003-066x.49.4.294

Language and Reading Research Consortium, Currie, N. K., & Muijselaar, M. M. (2019). Inference making in young children: The concurrent and longitudinal contributions of verbal working memory and vocabulary. *Journal of Educational Psychology*, *111*, 1416–1431. https://doi.org/10.1037/edu0000342

Language and Reading Research Consortium, Jiang, H., & Farquharson, K. (2018). Are working memory and behavioral attention equally important for both reading and listening comprehension? A developmental comparison. *Reading and Writing*, *31*, 1449–1477. https://doi.org/10.1007/s11145-018-9840-y

Lepola, J., Lynch, J., Laakkonen, E., Silvén, M., & Niemi, P. (2012). The role of inference making and other language skills in the development of narrative listening comprehension in 4–6-year-old children. *Reading Research Quarterly*, *47*, 259–282. https://doi.org/10.1002/rrq.020

Manly, T., Robertson, I. H., Anderson, V., & Nimmo-Smith I. (1998). *Test of Everyday Attention for Children* [Assessment instrument]. London, UK: Pearson.

Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first-and second-language learners. *Psychological Bulletin*, *140*, 409–433. https://doi.org/10.1037/a0033890

Moore, D. S., Notz, W. I, & Flinger, M. A. (2013). *The basic practice of statistics* (6th ed.). New York, NY: W. H. Freeman and Company.

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology*, *40*, 665–681. https://doi.org/10.1037/0012-1649.40.5.665

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, *98*, 554–556. https://doi.org/10.1037/0022-0663.98.3.554

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. https://doi.org/10.1075/lab.1.3.01par

Paris, A. H., & Paris, S. G. (2003). Assessing narrative comprehension in young children. *Reading Research Quarterly*, *38*, 36–76. https://doi.org/10.1598/rrq.38.1.3

Pearson, B. Z. (2007). Social factors in childhood bilingualism in the United States. *Applied Psycholinguistics*, *28*, 399–410. https://doi.org/10.1017/s014271640707021x

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*, 357–383. https://doi.org/10.1080/10888430701530730

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Vehoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–213). Amsterdam, The Netherlands: John Benjamins.

Proctor, C. P., Silverman, R. D., Harring, J. R., & Montecillo, C. (2012). The role of vocabulary depth in predicting reading comprehension among English monolingual and Spanish–English bilingual children in elementary school. *Reading and Writing*, *25*, 1635–1664. https://doi.org/10.1007/s11145-011-9336-5

R Core Team (2021). *R: A language and environment for statistical computing* (Version 4.1.1) [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Raudszus, H., Segers, E., & Verhoeven, L. (2021). Patterns and predictors of reading comprehension growth in first and second language readers. *Journal of Research in Reading*, *44*, 400–417. https://doi.org/10.1111/1467-9817.12347

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Ruffman, T. (1996). Reassessing children's comprehension-monitoring skills. In C. Cornoldi & J. Oakhill (Eds.) *Reading comprehension difficulties: Processes and intervention* (pp. 33–67). Mahwah, NJ: Lawrence Erlbaum.

Schönpflug, U., & Küpping-Faturikova, L. (2020). Bilingual children's story comprehension across language boundaries: Does retelling improve comprehension? *International Journal of Bilingual Education and Bilingualism*, *23*, 657–675. https://doi.org/10.1080/13670050.2018.1479727

Semel, E., Wiig, E., & Secord, W. (2006). *Clinical Evaluation of Language Fundamentals* (4th ed. UK) [Assessment instrument]. London, UK: Pearson.

Serratrice, L., & Hervé, C. (2015). Referential expressions in bilingual acquisition. In L. Serratrice, & S. Allen (Eds.), *The acquisition of reference* (pp. 311–333). Amsterdam, The Netherlands: John Benjamins.

Serratrice, L., & De Cat, C. (2020). Individual differences in the production of referential expressions: The effect of language proficiency, language exposure and executive function in bilingual and monolingual children. *Bilingualism: Language and Cognition*, *23*, 371–386. https://doi.org/10.1017/s1366728918000962

Silva, M., & Cain, K. (2015). The relations between lower and higher level comprehension skills and their role in prediction of early reading comprehension. *Journal of Educational Psychology*, *107*, 321–331. https://doi.org/10.1037/a0037769

Strasser, K., & Del Río, F. D. (2014). The role of comprehension monitoring, theory of mind, and vocabulary depth in predicting story comprehension and recall of kindergarten children. *Reading Research Quarterly*, *49*, 169–187. https://doi.org/10.1002/rrq.68

Sun, H., Yin, B., Amsah, F., & O'Brien, B. A. (2018). Differential effects of internal and external factors in early bilingual vocabulary learning: the case of Singapore. *Applied Psycholinguistics*, *39*, 383–411. https://doi.org/10.1017/S014271641700039X

Thordardottir, E. (2019). Amount trumps timing in bilingual vocabulary acquisition: Effects of input in simultaneous and sequential school-age bilinguals. *International Journal of Bilingualism*, *23*, 236–255. https://doi.org/10.1177/1367006917722418

Unsworth, S. (2016). Quantity and quality of language input in bilingual language development. In E. Nicoladis & S. Montanari (Eds.), *Bilingualism across the lifespan: Factors moderating language proficiency, Language and the Human Lifespan* (p. 103–121). Berlin, Germany: De Gruyter Mouton. https://doi.org/10.1515/9783110341249-008

Valentini, A., & Serratrice, L. (2021). What can bilingual children tell us about the developmental relationship between vocabulary and grammar? *Cognitive Science*, *45*(11), e13062. https://doi.org/10.1111/cogs.13062

Valentini, A., & Serratrice, L. (2022a). *Questionnaire. Materials from "Longitudinal predictors of listening comprehension in bilingual primary school-aged children"*

[Questionnaire]. IRIS Database, University of York, UK. https://doi.org/10.48316/zmv5-0863

Valentini, A., & Serratrice, L. (2022b). *Study materials. Materials from "Longitudinal predictors of listening comprehension in bilingual primary school-aged children"* [Collection: language task]. IRIS Database, University of York, UK. https://doi.org/10.48316/4vrt-a513

Valentini, A., & Serratrice, L. (2022c). *Data. Datasets from "Longitudinal predictors of listening comprehension in bilingual primary school-aged children*." IRIS Database, University of York, UK. https://doi.org/10.48316/3xdy-df60 </Dataset>

Wechsler, D. (2013). *Wechsler Preschool and Primary Scale of Intelligence* (4th UK ed.) [Assessment instrument]. London, UK: Pearson.

Whiteside, K. E., & Norbury, C. F. (2017). The persistence and functional impact of English language difficulties experienced by children learning English as an additional language and monolingual peers. *Journal of Speech, Language, and Hearing Research*, *60*, 2014–2030. https://doi.org/10.1044/2017_jslhr-l-16-0318

Wiig, E. H., & Secord, W. (1992). *Test of Word Knowledge* [Assessment instrument]. San Antonio, TX: The Psychological Corporation.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1: Principal Component Analysis Details for Memory and Vocabulary Depth.
Appendix S2: Backward Word Recall Task.
Appendix S3: Comprehension Monitoring Task.
Appendix S4: Listening Comprehension Task: Details of Piloting and Question Selection.
Appendix S5: Listening Comprehension Task: Stories.
Appendix S6: Language Questionnaire.

Appendix: Accessible Summary (also publicly available at https://oasis-database.org)

What predicts story comprehension in bilingual children?

What the Research Was About and Why It Is Important

Many children in the United Kingdom—as in many other countries around the world—can now be classified as bilinguals (i.e., they understand and/or speak more than one language). Just as it is for monolingual children (i.e., children who know only one language), understanding spoken language is an essential aspect of their language development. To make sense of what others say to them, children need to comprehend what is explicitly mentioned (literal understanding), and they often need to make informed guesses about what is being implied but not said, in other words, they need to make inferences. Literal comprehension and inferential comprehension are crucial for understanding stories, a ubiquitous experience for most children in and out of school.

Although some evidence exists for what predicts good story comprehension in monolingual children, relatively little is known about how bilingual children's cognitive skills (memory and attention) and their linguistic skills (knowledge of words and grammar) predict their literal and inferential comprehension of stories in their school language. Because bilingual children's language input is divided across two languages, another important predictor of how well they comprehend stories will be the quantity and length of exposure that they have had to the language of schooling, which was English in the context of our study.

What the Researchers Did

- At the beginning of Year 1 of primary school 100 bilingual children between the age of 5 and 6 years completed different tasks to assess their attention and memory skills, their word knowledge (vocabulary), and their grammar knowledge.
- At the beginning of Year 1, at the end of Year 1, and at the beginning of Year 2, children also listened to three stories (different ones each time) and answered questions on their understanding of these stories. The number of questions correctly answered formed their listening comprehension score. Questions were distinguished between literal questions (i.e., questions regarding events described in the story), local inference questions (i.e., questions where the children had to make connections between various aspects of the story), and global inference questions (i.e., questions where children had to apply their world knowledge to the story).
- During Year 1 parents completed a questionnaire about which languages were spoken in their homes and how often they were spoken. From this questionnaire, we derived a measure of English input (i.e., how much English the children had been exposed to up to that point).

What the Researchers Found

- Over time, children's story comprehension increased, especially for literal questions and global inference questions.

- Vocabulary depth (i.e., how much children know about words, measured through a synonym and opposite task) and morphological knowledge (i.e., children's ability to use the right morphological ending for the words, e.g., using *–s* for plurals or *–ed* for past tense) predicted their English language comprehension better than any other skill.
- English input influenced listening comprehension only indirectly by affecting vocabulary depth and morphological knowledge, which, in turn, affected story comprehension.

Things to Consider

- The results showed that the amount of information that children knew about words and their ability to use morphological endings correctly had the largest effect on their ability to understand spoken English in the context of a story where they needed to remember literal information and make inferences. This might suggest that the best way to improve bilingual children's spoken understanding of English would be to help them develop deep and rich vocabularies as well as teaching them morphology.
- Speaking more or less English at home was found not to affect children's literal and inferential understanding of English directly, but it did affect children's word knowledge and their grammar knowledge, which, in turn, affected their story comprehension.

Materials, data, open access article: Materials and data are publicly available on IRIS (www.iris-database.org) and OSF (https://osf.io/2fa3c/?view_only=cc167f7a3484432b9d5d56e4b87d4032).

How to cite this summary: Valentini, A., & Serratrice, L. (2022). What predicts story comprehension in bilingual children? *OASIS summary* of Valentini & Serratrice (2023) in *Language Learning*. https://oasis-database.org