Applied Psycholinguistics, page 1 of 28, 2012
doi:10.1017/S0142716411000865

Language abilities of internationally adopted children from China during the early school years: Evidence for early age effects?

AUDREY DELCENSERIE, FRED GENESEE, and KARINE GAUTHIER
McGill University

Received: July 29, 2010 Accepted for publication: July 22, 2011

ADDRESS FOR CORRESPONDENCE
Audrey Delcenserie, Department of Psychology, Stewart Biology Building, 1205, Dr. Penfield Avenue, McGill University, Montreal, QC H3A 1B1, Canada. E-mail: audrey.delcenserie@mail.mcgill.ca

ABSTRACT
We assessed the language, cognitive, and socioemotional abilities of 27 internationally adopted children from China, adopted by French-speaking parents, 12 of whom had been assessed previously by Gauthier and Genesee. The children were on average 7 years, 10 months old and were matched to nonadopted monolingual French-speaking children on age, gender, and socioeconomic status. Although there were no significant differences between the groups with respect to socioemotional and cognitive development, the adoptees scored significantly lower than the controls on measures of receptive grammar, expressive vocabulary, word definitions, and sentence recall, findings that were similar to those reported by Gauthier and Genesee. Analyses of correlations between the adopted children's language test results and their age at adoption, length of exposure to the adoption language, health, and other developmental problems revealed relatively few significant associations. In contrast, analyses of the relationship between their language test scores and their performance on the recalling sentences subtest suggest a link between performance on these two tests. We speculate on the role that performance on sentence recall might play in mediating differences in language outcomes between the two groups of children.

The primary goal of the present study was to examine the language development of school-age internationally adopted (IA) children from China during the early school years, including a group of IA children who had previously been evaluated by Gauthier and Genesee (2011). Thus, this study is the third phase of a longitudinal evaluation of IA children's language development spanning the preschool and early school age years. The language acquisition of IA children is unique in that they are exposed to a first language (L1) during several months, sometimes years, and then their exposure to this language is abruptly interrupted when they begin acquisition of their adoption language. Because of these unusual circumstances, along with other factors, the language development of IA children is often thought

to be at risk. First, IA children often experience preadoptive environments in the homes of their birth parents or in orphanages that may not meet their basic needs (e.g., nutritional) and that may even involve abuse, neglect, and inconsistent caretaking (Meacham, 2006), all conditions that can cause short- and long-term developmental delays (Glennen, 2002). Most orphanages have low caregiver–child ratios limiting the duration and frequency of interactions between IA children and a single caregiver, which may lead, in turn, to a lack of physical, cognitive, and social stimulation. Evidence suggests that the duration of institutionalization, along with the severity of deprivation preadoptively, might explain the poor speech and language development exhibited by some adoptees and, in particular, those from Eastern Europe (Glennen, 2002; Groze & Ileana, 1996; Meacham, 2006). The extent to which their preadoptive environment puts some IA children at risk may be related to the country of adoption, given that the reasons for abandonment and the quality of institutional care adopted children receive vary among countries (Gunnar, Bruce, & Grotevant, 2000; Hyltenstam, Bylund, Abrahamsson, & Park, 2009).

Second, many, although not all, IA children discontinue exposure to and acquisition of the birth language upon adoption and this abrupt termination of L1 acquisition may influence their acquisition of the adoption language. According to the exercise hypothesis, the capacity for language learning must be exercised early in life so that it remains intact for subsequent language acquisition (Johnson & Newport, 1989). All of the IA children in the present study discontinued exposure to and acquisition of their birth language, and this raises the question of whether discontinuing acquisition of the L1 undermines the neurocognitive substrates for acquisition of the adoption language (e.g., Johnson & Newport, 1989).

Third, another related factor that could compromise IA children's acquisition of their new language is their delayed onset of exposure to the adoption language. According to the classic version of the critical period hypothesis, language learning is more effective and complete the earlier it begins and is more likely to result in less than nativelike competence as age of onset is delayed (Penfield & Roberts, 1959). Under this hypothesis, one would expect IA children who are adopted within 1 to 2 years of birth to acquire full, nativelike competence in their adoption language. However, there is recent evidence for very early age effects on second-language (L2) learning, much earlier than previously thought (e.g., Hyltenstam, 1992). More specifically, Abrahamsson and Hyltenstam (2009) found that, despite more than 20 years of exposure to Swedish as an L2, only 3 of the 31 preschool learners of Swedish (i.e., 1–5 years of age) they assessed performed like native Swedish speakers on all measures of an extensive battery of language ability and processing tasks they administered (e.g., cloze test, speech perception in noise). It might thus be expected that IA children who acquire a new language 1 to 2 years after birth would exhibit similar early age effects (see also Hyltenstam et al., 2009). In the case of many IA children, delay in exposure to the adoption language is confounded with disruption in acquisition of the birth language, a point we return to later.

Although IA children's language development may be at risk for several reasons, there are other reasons for expecting that they might be advantaged in learning their new language in comparison to other L2 learners. First, adoptive parents have

higher than average levels of income and education (e.g., Hellerstedt et al., 2008; Roberts et al., 2005), and these factors would be expected to have positive effects on the amount and type of parent talk and the quality of interaction IA children experience postadoption (Tan & Yang, 2005). Mothers with high education levels and socioeconomic status (SES) tend to speak more to their children than mothers from lower SES backgrounds, and their children in turn have been found to exhibit above average expressive and receptive vocabulary, general language abilities, and more lexically complex utterances (e.g., Hoff, 2003, 2006; Hoff & Tian, 2005; LeNormand, Parisse, & Cohen, 2008).

Second, despite much controversy concerning the existence of a critical period for language development and the precise termination of this period, there is still considerable evidence that early L2 learning is more successful than L2 learning that occurs later in life (e.g., Birdsong & Molis, 2001; DeKeyser, 2000; Johnson & Newport, 1989). Thus, because IA children from China are adopted early, usually before 24 months of age, one might expect no adverse effect of age of acquisition of their new language. In addition, because they are exposed to the new language so early, the neurocognitive mechanisms that underlie L1 acquisition might still be fully available for acquisition of the adoption language.

Third, in contrast to successive and simultaneous bilinguals who are exposed to and learn two languages, most IA children's exposure to the adoption language is not divided between an L1 and an L2. They, therefore, do not need to acquire or process more than one language at a time because they benefit from full exposure to their new language and this should in turn facilitate acquisition of that language.

LANGUAGE DEVELOPMENT OF IA CHILDREN

Research has shown that despite risk factors, most preschool IA children perform within the normal range when assessed using parent reports or standardized tests designed for monolingual speakers of the same age (Geren, Snedeker, & Ax, 2005). Indeed, based on parental reports of vocabulary development, such as the MacArthur Communicative Developmental Inventory (Glennen, 2002), Chinese adoptees have been found to perform at the same level as native English speakers within 12 months postadoption (Snedeker, Geren, & Shafto, 2007). In the remainder of this report, and for ease of reference, we refer to native speaking children of the adoption language (e.g., English in the United States) as "nonadopted children," although we recognize that there are children in orphanages in China and elsewhere who are also not adopted. It has also been found that more than 85% of children who were adopted from China between 6 and 25 months of age and living in English-speaking homes performed within or above the average range on other standardized language measures, such as the Peabody Picture Vocabulary Test, at 30–47 months postadoption (Roberts & Krakow, 2003). Evidence shows that, in general, IA children who are adopted before 24 months of age often achieve nativelike levels of proficiency within 12 months postadoption, thereby demonstrating a rate of language acquisition that is faster than that of native speakers of the target language. For example, Snedeker et al. (2007) found that preschool Chinese adoptees, who were between 2 years, 7 months (2;7) and 5;6 at the time of testing and assessed every 3 months had a vocabulary size that was

similar to that of native English-speaking children of 24 months of age after only 3 months of exposure to English.

Although IA children make impressive gains in acquiring their new language postadoption, their language outcomes depend on their age at adoption and their length of exposure to their new language. There is well-documented evidence that children adopted at younger ages, typically before 12 months of age, display better language outcomes during the preschool years and attain nativelike language proficiency sooner than children adopted at older ages (e.g., Krakow, Tao, & Roberts, 2005). In contrast, older IA children have been shown to exhibit faster rates of acquisition initially, and in the short term, but also to be less likely to achieve parity with native speakers in the long term (e.g., Glennen, 2009). For example, Pollock (2005) assessed the vocabulary growth of children who were adopted from China at different ages and found that children who were adopted after 24 months of age had a vocabulary of 400 words after 6 months of exposure to English, whereas children adopted before 12 months of age had a vocabulary of 50 words. Although the older adoptees exhibited faster initial vocabulary development than the children adopted at younger ages, as just noted, the children adopted at older ages had more to learn to catch up to same-age native speakers and, therefore, required more time to score within the typical range for their age.

Another factor that can influence IA children's language abilities is their preadoptive language environment. The primary reason why children in the present study were given up for adoption in China is this country's strict birth planning policy, a one-child policy designed to reduce population growth. Most parents who abandon children in China, usually girls, are married, of average SES, and come from rural areas, creating a preadoptive environment that is more advantageous than in other countries (e.g., Gauthier & Genesee, 2011; Johnson, Banghan, & Liyao, 1998). In contrast to IA children from other countries, Chinese adoptees are therefore less likely to suffer from the effects of parental alcoholism, drug abuse, and/or poor mental health, poverty, general neglect and abuse, and familial dysfunctionality (Paradis, Genesee, & Crago, 2011).

In a related vein, the language outcomes of adoptees can also be related to their country of origin. All of the adopted children in the present study had been in orphanages at the time of adoption, according to parent reports. Although Chinese orphanages often provided suboptimal care for abandoned children in the past, evidence shows that the situation has improved (e.g., Hwa-Froelich & Matsuoh, 2008; Johnson et al., 1998). Accordingly, IA children from China are healthier and thus less prone to exhibit language difficulties that are associated with early health problems in comparison to children adopted from other countries (e.g., Cohen, Lojkasek, Zadeh, Pugliese, & Kiefer, 2008; Roberts, Krakow & Pollock, 2003). In addition, parents usually make the decision to put their children up for adoption early, within the first 6 months after birth (Johnson et al., 1998), so that adoptees from China are adopted relatively early and are consequently institutionalized for shorter periods of time. This might explain why they have fewer health and development difficulties as well as fewer socioemotional and cognitive problems, all factors that can affect language development.

Notwithstanding evidence of normal development for many IA children during the preschool years, there is also evidence of difficulties in a subgroup of IA

children that is larger than what one finds in children in the population at large (e.g., Roberts et al., 2003). To be more specific, evidence shows that IA children exhibit higher referral rates for assessment and treatment by speech–language pathologists than nonadopted children. For example, in a longitudinal study by Glennen and Masters (2002) in which they collected data on Eastern European adoptees' language development using surveys and language scales, it was found that, of the 130 children who had been adopted from below 6 months of age to 30 months of age, 53.8% had speech–language assessments, whereas 64.3% were recommended for speech–language therapy. Moreover, studies that have directly assessed IA children's language development during preschool have also found evidence of difficulties. For example, and of particular relevance to the present study, is the longitudinal study by Gauthier and Genesee (2011) that examined the language, cognitive, and socioemotional development of Chinese adoptees. The children were an average of 1;2 at the time of adoption and were acquiring French; they were assessed twice, once between 3;6 and 4;8 and again between 4;9 and 6;0. Their language abilities were compared to that of nonadopted monolingual French-speaking children matched for gender, age, and SES. Their results showed that the expressive language and vocabulary scores of the Chinese adoptees were significantly lower than those of native French-speaking adopted children as were their receptive language skills and their performance on the recalling sentences subtest of the Clinical Evaluation of Language Fundamentals—Revised (see also Cohen et al., 2008). Gauthier and Genesee argued that their results were not due to the children's preadoptive learning environment because they performed, as a group, in the normal range on most measures in all domains (socioemotional, cognitive, and language). They argued further that their results were not due to amount of exposure. Were exposure the primary factor, one would have expected the gap between the IA and comparison children to have been reduced at the second assessment, and they did not find this. However, because the comparison children had also had more exposure to French, it could be argued that the exposure that the IA children had had to French at the time of their second assessment was insufficient. Thus, the present study was undertaken to examine this possibility.

The question remains, nevertheless, whether the IA children examined by Gauthier and Genesee (2011) would reach parity with nonadopted CTL children with yet more exposure to the adoption language, especially if that additional exposure occurred in the context of schooling. Schooling is an enriched and challenging learning environment in which children are taught to extend their language competencies for abstract, cognitively demanding, and complex communication. On the one hand, it might be expected that this language learning environment would provide enrichment that would enhance the language abilities of IA children and in particular the minority of adoptees who exhibit lags in development during the preschool years. In contrast, the linguistic demands of schooling might challenge IA children's language abilities further, resulting in continued or possibly even greater lags in their language abilities relative to nonadopted peers.

The results of research on the language, academic, and cognitive development of IA children during the early school years are mixed. Generally speaking, the majority of IA children demonstrate considerable resilience in cognitive, academic, and linguistic development during the school years. A majority of IA children

perform similarly to their classmates, environmental siblings, or peers in the general population on measures of cognitive ability, such as IQ (van IJzendoorn, Juffer, & Poelhuis, 2005), and on measures of academic ability, such as parent and/or teacher reports of school performance and grades (Dalen & Ryvgold, 2006). Studies have also reported that IA children demonstrate relatively good language outcomes during the school years on several measures, including tests of receptive vocabulary, reading, writing, narrative abilities, and everyday language (Andresen, 1992; Clark & Hanisee, 1982; Croft et al., 2007; Scott, Roberts, & Krakow, 2008).

There is also considerable evidence that more school-age IA children experience language and academic difficulties than is found in the general population of school children. Evidence from rates of referral for speech–language assessment and treatment indicates that a substantial proportion of IA children may exhibit delays or difficulties in comparison to the general population of nonadopted children, and also that these referral rates increase with age at adoption. For example, in a longitudinal study of IA children from Eastern European countries, Glennen and Masters (2002) found that 47% of IA children who were adopted before 12 months of age, 58% of children who were adopted between 13 and 18 months of age, and 73% of children who were adopted between 19 and 24 months of age were referred for speech–language assessments. Because referral rates for speech–language assessments can reflect adoptive parents' level of concerns for the language development of their child, these results should be interpreted with caution.

There is additional evidence from direct assessments that the language difficulties of IA children during the early school years can persist even with additional exposure to the adoption language. For example, Roberts, Pollock, and Krakow (2005) monitored the language development of 10 low-performing IA children who had been identified in an earlier investigation (see Roberts et al., 2003). The children were an average of 5;10 at the time of the follow-up assessment, which was some 2;3 after the initial assessment. The purpose of the study was to ascertain whether the additional exposure to English that occurred between the initial and follow-up assessment would reduce or close the gap with the norming group. Their language abilities were examined using a battery of standardized language measures that assessed their expressive and receptive vocabulary, language abilities, and articulation. Results showed that, although the adoptees made considerable gains on these measures from the initial to the follow-up assessment, their performance continued to be significantly lower than that of the comparison group despite two additional years of exposure to their adoption language.

It can be difficult to interpret results from these studies because different studies and even sometimes the same studies included children with different ages at adoption and/or from different countries. Studies that include a relatively high proportion of IA children who were relatively old when adopted and/or from countries with institutionalized care that is seriously impoverished could skew the results of the entire sample to the low end of performance, given the association between these two factors and language outcomes discussed earlier.

In a meta-analysis of 62 studies ($N = 17{,}767$) of school-age adopted children that could potentially control for these confounding factors, van IJzendoorn et al. (2005) reported no significant differences between adopted children, both domestic and international, and nonadopted children on measures of cognitive

ability, but that adopted children were rated significantly lower, albeit the differences were small, on parent and teacher ratings of language and school achievement compared to nonadopted peers in the same community as the IA children. However, children from China were not included in this analysis and there were no direct assessments of the children's language or academic performance.

THE PRESENT STUDY

The present study sought to examine the long-term language outcomes of IA children from China into the school years. In fact, the study was a partial extension of Gauthier and Genesee's (2011) study insofar as a subsample of the adoptees in the present study ($n = 12$) were children who had participated in Gauthier and Genesee's (in press) study. This permitted us to conduct a longitudinal assessment of their language development. An additional 15 IA children were added to increase the sample size; this permitted us to examine the generalizability of results from Gauthier and Genesee's subsample. The primary objective was to determine if lags in language development, as exhibited by the preschool IA children examined by Gauthier and Genesee, were still evident in older IA children. Evidence that earlier lags were resolved in the present study would argue that the children examined by Gauthier and Genesee simply had not had sufficient exposure to master all aspects of French. In contrast, evidence of a persistent lag in the present study would argue that other factors are at play, possibly early age effects.

In contrast to previous studies of school-age IA children that have tended to rely on parent/teacher report measures or comparisons with test norms, the present study included a comparison group of nonadopted children who were carefully matched with the IA children on age, socioeconomic status, and gender. All of these factors can influence language learning, to varying degrees, but are seldom taken into account in other studies (in contrast, see Cohen et al., 2008; Gauthier & Genesee, 2011). As a result, extant research does not necessarily provide a complete picture of the language development of school-age IA children relative to same age peers when factors that are known to influence language development are taken into account, in particular, socioeconomic status. As demonstrated by Gauthier and Genesee, as well as Cohen and her colleagues, when direct comparisons are made between IA and carefully matched comparison groups, a more differentiated profile of similarities and differences emerges, with evidence of relatively low performance for IA children, than when only test norms or report measures are used.

METHOD

Participants

As noted earlier, it was possible to recruit only 12 of the 24 IA children tested by Gauthier and Genesee (2011) to participate in the present study. The major reason parents reported for nonparticipation was lack of time. Therefore, an additional 15 girls adopted from China by French-speaking parents were recruited to increase the sample size to 27. The children had been adopted between 7 months

and 1;9 ($M = 12.9$ months, $SD = 3.8$ months). As a group, they had had a mean length of exposure to French of 6;9 ($SD = 7.4$ months), and were between 7;0 and 8;8 ($M = 7;10$, $SD = 6.0$ months) at the time of testing. Families of the adopted children were invited to participate using either contact information we already had for returning children or with the assistance of adoption agencies for new recruits.

The IA children were compared to a control group of 27 monolingual non-adopted French-speaking children who were between 6;9 and 8;10 at the time of testing ($M = 7;11$, $SD = 6.9$ months). Participants were in Grades 1 to 3. The control (CTL) children were recruited from local schools and were matched to the IA children to within 6 months of birth and on SES (see Appendix A). Exclusionary criteria for controls were (a) presence of psychiatric or neurological antecedents; (b) history of intellectual deficiency and language problems; (c) premature birth; (d) serious health, motor, or behavior problems; (e) L1 other than French; and (f) more than 25% exposure to an L2, as reported by parents on the Language Environment Questionnaire, described in the next section.

Questionnaires and assessment materials

The Developmental Questionnaire, which was used by Gauthier and Genesee and adapted from the Language Development Questionnaire for Children Adopted from Eastern Europe (Glennen & Masters, 2002), was given to parents to collect information about each child's health, behavior, development, socioemotional adjustment and about parents' age, level of education, and income.

The Child Behavior Checklist (CBCL; Achenbach, 1991a), a parent report designed for children between 6;0 and 12;0, was used to assess the children's behavior and social competence at home. Twenty-four adoptive parents and 24 control parents completed the questionnaire. The major reason reported by parents for not completing the questionnaire was lack of time. Close examination of the results of the six children whose parents did not complete the CBCL revealed that these children scored in the average or above average range on the language tests in comparison to the other adoptees.

The CBCL—Teacher Report Form (CBCL-TRF; Achenbach, 1991b), which is similarly appropriate for children between 6;0 and 12;0, was used to elicit teacher's perceptions of the children's academic performance, adaptive functioning, and behavior problems in school. Eighteen questionnaires were completed by IA children's teachers, and 22 were completed by the teachers of CTL children.

The Language Environment Questionnaire from Gauthier and Genesee (2011) was used to ascertain the amount of input children received in French or other languages from all family members and in diverse situations (i.e., television, at school).

The Wechsler Non-Verbal IQ test (Wechsler & Naglieri, 2006) was used to assess general cognitive abilities. This test was chosen because it can be administered and completed without the use of language, making it suitable for use with French-speaking children; as well, the influence of language ability on test performance is minimized. The test is appropriate for children between 4;0 and 21;11, although different subtests are recommended for children between 7 and 8 years of age. IA

and CTL children between 7;0 and 7;11 were administered the matrices, coding, object assembly, and recognition subtests, and children between 8;0 and 8;11 were administered the matrices, coding, spatial span, and picture arrangement subtests. Children who were above 7;11 were administered the tests for 8-year-olds. For the matrices subtest, the children examined an incomplete geometric figure and selected the missing portion from five response options. For the coding subtest, the children copied symbols paired with simple geometric shapes or numbers. Using a key, the children had to copy a series of symbols that corresponded to a series of shapes within a specific time limit. For the object assembly subtest, the children were presented with prearranged puzzle pieces and had to fit the pieces together to form a meaningful whole (e.g., an apple, a bear) within a specified time limit. For the recognition subtest, the children inspected a series of complex geometric designs for 3 s each and then identified which of four or five options matched the target shape. The options and the target differed, sometimes subtly, in terms of colors and patterns. For the spatial span subtest, children had to tap a series of blocks, forward or backward, according to a sequence demonstrated by the examiner. For the picture arrangement subtest, the children were asked to reorder sets of picture cards to tell a logical story within a specified time limit.

The Expressive One Word Picture Vocabulary Test—Third Edition (EOWPVT; Brownell, 2000, French adaptation) was used to assess expressive vocabulary. This test, a French adaptation of the original English version, assesses children's ability to make word–picture associations. Children were shown pictures that they then had to name. Although the psychometric properties of this version of the test are not comparable to those of the English version, this adaptation is used widely by the Speech and Language Pathology Department of the Montreal Children's Hospital and was used in previous assessments by Gauthier and Genesee. The results for five IA children were omitted due to irregularities in testing procedures.

The Échelle de Vocabulaire en Image Peabody (EVIP; Dunn, Theriault-Whalen, & Dunn, 1993) was used to assess receptive vocabulary skills and had been used in Gauthier and Genesee's previous assessments. In this test, the child has to find the image from among a set of four images that corresponds to a word spoken by the tester. Norms for French-speaking Canadians are available for this test. The internal validity of this test, as measured by the Claparède Indice (Gauthier & Genesee, 2011), indicates that the EVIP is a sensitive test up to 13 years of age, after which the results are more representative of individual differences related to age. This test is thus sensitive for the age range of the children.

The French version of the reading comprehension subtest of the Wechsler Individual Achievement Test—Second Edition (WIAT-II; Wechsler, 2005) was used to assess reading comprehension skills, a school-specific language ability. The children were required to read single sentences and short texts of 7 to 10 sentences in length and then to answer a series of questions about the content of each text. Norms for French-speaking Canadians are available.

The word definition subtest of the Wechsler Intelligence Scale for Children—Fourth Edition (Wechsler, 2003), French version, was used to assess the children's expressive language skills. This subtest requires children to correctly recall and coherently express the definition of words (e.g., éponge, île). Norms for French-speaking children are available.

The word associations and recalling sentences subtests of the Clinical Evaluation of Language Fundamentals—Revised (Semel, Wiig, & Secord, 1987), French adaptation, were used to assess semantic category knowledge and verbal memory abilities, respectively. The word associations subtest assesses children's early literacy abilities by requiring them to name words in specific semantic categories (e.g., food, animals, and professions) within 1 min. For the recalling sentences subtest, the children were asked to repeat sentences, varying in length and syntactic complexity, presented by the experimenter. This subtest was included because Gauthier and Genesee found that the IA children in their study scored significantly lower than the CTL children on this test and also significantly below test norms. In fact, they scored more than 1 *SD* below the norm. This was the only test on which their performance was below test norms. This is noteworthy because it has been found that tests of sentence recall are sensitive to age of acquisition among both first and second language learners (Mayberry & Eichen, 1991; Mayberry & Fischer, 1989) and performance on such tests is often used as a clinical marker of specific language impairment (e.g., Conti-Ramsden, Botting, & Faragher, 2001; Stokes, Wong, Fletcher, & Leonard, 2006). Although Gauthier and Genesee were careful to emphasize that the IA children were not language impaired in the clinical sense, they proposed that their performance on this test indicates that they had difficulty with aspects of French that show age sensitivity and are difficult for children with SLI. In other words, IA children's acquisition of French is vulnerable in the same way as some researchers have suggested is the case for other learners of French.

Sentence recall is thought to reflect children's phonological short-term memory abilities and, more specifically, the capacity of their phonological short-term store (Baddeley, 2003), as well as components of long-term memory (Alloway & Gathercole, 2005), probably linked to knowledge of syntax because children have to comprehend the syntactic constructions of the sentences in order to memorize and repeat them correctly (Semel et al., 1987). There is evidence that phonological memory improves with language development (French & O'Brien, 2008) and, in particular, that the ability to actively rehearse the content of the phonological store matures around 7 years of age allowing children to hold more material in memory without decay (Gathercole, n.d.). Including the recalling sentences subtest allowed us to examine if IA children's ability on this test would improve with age and more exposure to the adoption language or whether they would continue to perform below that of nonadopted children and age norms.

We also sought to determine whether the lags in memory demonstrated by the IA children on the recalling sentences subtest would extend to tests that assess visuospatial short-term memory. Accordingly, we administered the recognition and the spatial span subtests of the Wechsler Non-Verbal IQ test. Both assess visuospatial short-term memory and, more precisely, the capacity of children's visuospatial sketchpad (Baddeley, 2003). Including these tests permitted us to ascertain whether the IA children's lags in verbal memory relative to controls are similar to those of children with low working memory, who typically have difficulties on measures of both verbal and visuospatial short-term memory, or if, in contrast, their memory difficulties are specific to verbal memory (Alloway, Gathercole, Kirkwood & Elliott, 2009).

The Épreuve de Compréhension Syntaxico-Sémantique (ECOSSE; Lecocq, 1996) is a French version of the Test for Reception of Grammar (Bishop, 1983) and was used to assess children's receptive language abilities and, more specifically, their comprehension of syntax. Knowledge of pronouns, adjectives, negative phrases, and word order were assessed. The experimenter read a sentence aloud to the child, such as "La vache pousse la dame" (*The cow pushes the woman*), and the child had to find the image that corresponded to the sentence from among a set of four images. Correct responding depended on understanding of target forms, such as passive voice and relative pronouns. Norms for French-speaking children are available.

All the standardized measures included in the present study have reliability coefficients that are at or above .80.

Procedure

Before testing began, the experimenter explained the procedure of the study to the parents and children, presented the questionnaires to the parents, and answered questions. Parents were then asked to sign the consent form. Parents who consented to participate were then asked to complete the CBCL, the Language Environment Questionnaire, and the Developmental Questionnaire. They were allowed to fill out these forms quietly in the room during testing; however, most preferred to fill them out between the first and second testing sessions, separated by no more than 15 days. Depending on parents' preferences, the CBCL-TRF was either given directly to the child's teacher by the experimenter or to parents who in turn gave it to their child's teacher.

Each child was tested individually in a separate room at the university, in the child's school, or in the home. When testing was carried out in the home or at the university, parents were asked to remain as quiet as possible and not to help their child. Each testing session lasted about 50 min, and two sessions were required. The order of administration of the tests was counterbalanced to avoid biases due to order effects. Narrative and clitic elicitation tasks were also administered during these testing sessions, but these results will not be presented in this report.

RESULTS

Demographic information

One-way independent groups analyses of variance (ANOVAs; $\alpha = 0.05$) were carried out to compare the IA and CTL children on age at testing, fathers' age, and mothers' age (see Appendix A). Results showed that the groups differed significantly only in terms of fathers' age, $F\,(1, 30) = 8.45$, $p = .01$, and mothers' age, $F\,(1, 31) = 11.13$, $p < .01$, with adoptive parents being significantly older than control parents. Chi-square tests ($\alpha = 0.05$) were performed to compare the IA and CTL children in terms of school grade, educational level of each parent, and family income. No significant differences were found between the groups in terms of school grade, $\chi^2\,(2, 54) = 2.88$, $p = .24$, educational level of mother,

χ^2 (2, 54) = 0.92, $p = .63$, educational level of father, χ^2 (2, 54) = 1.06, $p = .59$, and family income, χ^2 (2, 54) = 1.39, $p = .85$.

General health and socioemotional development

As in the case of Gauthier and Genesee's study, information was collected about each child's past and current health status. None of the parents in either the IA or CTL group expressed concerns about their child's current general health; see Appendix B for a summary of information about the children's past and present health status. With regards to the children's past health status, results revealed that the IA and CTL children had had a comparable number of reported problems. The reported problems for the CTL children were mostly related to ear infections, which are common in infants and young children (Hôpital Ste-Justine, 2009), whereas the main problems reported for the IA children by their parents were emotional and attachment difficulties, as well as behavioral problems. With respect to current health status, results revealed that although the IA children had fewer health problems than in the past, they currently had more problems than the CTL children. However, 38% of the reported problems of the IA children were related to vision and these were in no case severe; moreover, there is evidence showing that children of Asian descent have more vision difficulties than Hispanic or Caucasian children (Kleinstein et al., 2003). More IA children (11.1%) were diagnosed with attention-deficit/hyperactivity disorders than CTL children (0%); but the results for the IA children are still within the normal range for this age group (Glennen & Bright, 2005).

In terms of socioemotional development, the IA children's performance on the CBCL and the CBCL-TRF indicated that they were developing as well as their nonadopted peers. Indeed, one-way independent-groups ANOVAs ($\alpha = 0.05$) were performed to compare children's total scores, scores on the *internalizing* and *externalizing* scales of the questionnaires, and on each subscale composing these scales. Results showed that the scores of the IA and CTL children were not significantly different on any of these, except for the social problems subscale of the CBCL-TRF, F (1, 38) = 5.27, $p = .03$, on which the CTL children were reported to have more problems than the IA children. The latter results corroborate the findings of previous studies that IA children are generally well adjusted, and also that Chinese adoptees tend to obtain scores on standardized assessments of emotional and behavioral abilities, such as the CBCL, that are similar or better than the norms (e.g., Rojewski, Shapiro, & Shapiro, 2000; Tan & Marfo, 2006).

Academic and nonverbal cognitive development

In terms of cognitive development, the IA children's total raw scores on the Wechsler Non-Verbal IQ test as well as their raw scores on each subtest (the matrix, the coding, the recognition, the image arrangement, the spatial span, and the object assembly subtests) were compared to those of the CTL children using one-way independent-groups ANOVAs ($\alpha = 0.05$). No significant differences were found between the IA and CTL children on the total raw score, F (1, 52) = 2.63, $p = .11$,

or on the total standard score, $F\,(1, 52) = .10$, $p = .76$. However, the IA children had significantly higher scores on the coding subtest than the CTL children, $F\,(1, 52) = 5.45$, $p = .02$, suggesting above normal nonverbal processing speed for the IA children; no other significant differences were observed. Otherwise, there were no significant group differences on the other subscales of this test.

Comparisons between the IA and CTL children on academic performance, using teachers' responses on the adaptive functioning subscale of the CBCL-TRF (including the following scales: works hard, behaves appropriately, learning level, and happiness), revealed no significant difference between the groups. Although none of our participants repeated a grade, five IA children and five CTL children were reported by their parents and/or teachers, on the CBCL and the CBCL-TRF, to have academic abilities that were below those of their same-age peers. In contrast to some of the CTL children who were reported to have difficulties in French, English as an L2, mathematics, and history, the five IA children who were reported to have academic difficulties all had difficulties with French. As will become evident in the next section, our statistical analyses of the IA children's language results revealed difficulties on the part of the IA children in comparison to the CTL children with respect to French.

Language development

Parental responses to the Developmental Questionnaire revealed that, although approximately the same number of parents of IA and CTL children consulted specialists, speech–language pathologists were the specialists consulted most often by adoptees and their parents ($n = 6$). Our results also revealed that these same six IA children (22.2% of our sample) had had or were still receiving speech–language therapy. These results are in line with previous studies that have found, as mentioned earlier, that a larger than expected subgroup of IA children require the services of speech–language professionals (Glennen & Bright, 2005; Scott et al., 2008). It is interesting that the adoptees who were identified by their parents or teachers as having lower academic abilities than their nonadopted peers were not the same IA children who were receiving speech–language therapy.

In terms of language test results, the IA and CTL children were compared using one-way independent-groups ANOVAs ($\alpha = 0.05$; see Table 1). The performance of the IA children was significantly lower than that of the CTL children on the ECOSSE, the EOWPVT, the word definitions subtest, and the recalling sentences subtest. The groups did not differ significantly on the word association subtest, the EVIP, or the WIAT.

To better characterize the performance of the IA children, the distribution of their scores was compared to that of the CTL children by calculating the percentage of IA children who scored above and below the average of the CTL children in terms of standard deviations (see Table 2). IA children who scored between $+1$ and -1 *SD* of the mean of the CTL children were considered to have a "normal" score compared to their nonadopted peers. A significant percentage of IA children performed at least 1 *SD* below the mean of the CTL children. More specifically, 66.6% of the IA children scored at least 1 *SD* below the mean of the CTL children

Table 1. *Results of tests (raw scores)*

Measures	IA	CTL		
	M (*SD*)	*M* (*SD*)	*df*	*F*
CELF-R				
Recalling sentences subtest	36.37 (11.99)[a]	50.33 (9.60)[a]	(1, 52)	22.31**
Word association subtest	32.30 (9.91)[a]	34.07 (7.52)[a]	(1, 52)	0.55
ECOSSE	10.14 (4.37)[a]	7.30 (2.54)[a]	(1, 52)	8.61**
EOWPVT	78.36 (14.65)[b]	87.70 (10.27)[a]	(1, 47)	6.86**
EVIP	107.29 (12.12)[a]	109.48 (12.84)[a]	(1, 52)	0.41
WIAT-II				
Reading comprehension subtest	33.19 (5.47)[a]	33.74 (7.18)[a]	(1, 52)	0.10
WISC-IV				
Word definition subtest	20.62 (6.49)[a]	24.63 (4.42)[a]	(1, 52)	7.00**
Wechsler Non-Verbal IQ				
Matrix subtest	10.89 (2.53)[a]	11.44 (1.97)[a]	(1, 52)	0.81
Coding subtest	38.74 (10.42)[a]	32.93 (7.67)[a]	(1, 52)	5.45*
Object assembly subtest	41.72 (8.09)[c]	43.14 (5.72)[d]	(1, 30)	0.31
Recognition subtest	12.39 (2.45)[c]	13.14 (2.71)[d]	(1, 30)	0.68
Spatial span subtest	14.44 (4.53)[e]	13.61 (2.81)[f]	(1, 20)	0.28
Image arrangement subtest	5.11 (4.14)[e]	7.31 (4.57)[f]	(1, 20)	1.32
Total score	92.22 (19.52)[a]	83.63 (19.43)[a]	(1, 52)	2.63

Note: IA, internationally adopted; CTL, control; CELF-R, Clinical Evaluation of Language Fundamentals—Revised; ECOSSE, Épreuve de Compréhension Syntaxico-Sémantique; EOWPVT, Expressive One-Word Picture Vocabulary Test—Third Edition; EVIP, Échelle de Vocabulaire en Images Peabody; WIAT-II, Wechsler Individual Achievement Test—Second Edition; WISC-IV, Wechsler Individual Scale for Children—Fourth Edition. The raw scores for the ECOSSE represent the number of errors.
[a] $n = 27$.
[b] $n = 22$.
[c] $n = 18$.
[d] $n = 14$.
[e] $n = 9$.
[f] $n = 13$.
*$p < .05$. **$p < .01$.

on the recalling sentences subtest, 29.6% on the word association subtest, 48.1% on the ECOSSE, 50% on the EOWPVT, 22.2% on the EVIP, 11.1% on the WIAT, and 51.8% on the word definition subtest.

The preceding statistical analyses were performed using raw scores on the language tests. The IA children's performance was also compared to norms for French-speaking children, when these were available. When the norms for French-speaking children were not available, the norms for English-speaking Canadians were used, but must be interpreted with caution. Results indicated that the IA children, on average, performed at or above the norms on all language tests, except for the recalling sentences subtest. On the latter, the IA children's results

Table 2. *Percentage of internationally adopted children who scored above and below the mean of control children on language tests*

SD	Recalling Sentences (CELF-R)	Word Association (CELF-R)	ECOSSE	EOWPVT	EVIP	WIAT-II	WISC-IV
−2	29.6	7.4	22.2	18.2	0	0	22.2
−2 and −1.25	33.3	22.2	11.1	9.1	7.4	11.1	22.2
−1.25 and −1	3.7	3.7	14.8	22.7	14.8	0	7.4
−1 and 0	18.5	11.1	22.2	22.7	33.3	33.3	22.2
0	0	0	0	0	0	0	0
0 and +1	11.1	37.0	25.9	13.6	33.3	55.6	18.5
+1 and +1.25	3.7	7.4	0	13.6	3.7	0	0
+1.25 and +2	0	11.1	3.7	0	7.4	0	3.7
+2	0	0	0	0	0	0	3.7

Note: CELF-R, Clinical Evaluation of Language Fundamentals—Revised; ECOSSE, Épreuve de Compréhension Syntaxico-Sémantique; EOWPVT, Expressive One-Word Picture Vocabulary Test—Third Edition; EVIP, Échelle de Vocabulaire en Images Peabody; WIAT-II, Wechsler Individual Achievement Test—Second Edition; WISC-IV, Wechsler Individual Scale for Children—Fourth Edition.

Table 3. *Results of tests (standard scores)*

Measures	IA	CTL
	M (*SD*)	*M* (*SD*)
CELF-R		
Recalling sentences subtest	6.44 (2.64)[a]	9.41 (1.99)[a]
Word association subtest	12.56 (4.07)[a]	12.74 (2.81)[a]
ECOSSE	8.15 (1.79)[a]	9.30 (1.61)[a]
EOWPVT	99.09 (14.53)[b]	104.70 (21.02)[a]
EVIP	126.04 (11.50)[a]	126.26 (11.19)[a]
WIAT-II	61.00 (9.83)[a]	58.41 (10.54)[a]
WISC-IV		
Word definition subtest	8.48 (2.52)[a]	9.96 (2.38)[a]
Wechsler Non-Verbal IQ	95.67 (9.77)[a]	94.85 (9.32)[a]

Note: IA, internationally adopted; CTL, control; CELF-R, Clinical Evaluation of Language Fundamentals—Revised; ECOSSE, Épreuve de Compréhension Syntaxico-Sémantique; EOWPVT, Expressive One-Word Picture Vocabulary Test—Third Edition; EVIP, Échelle de Vocabulaire en Images Peabody; WIAT-II, Wechsler Individual Achievement Test—Second Edition; WISC-IV, Wechsler Individual Scale for Children—Fourth Edition. The raw scores for the ECOSSE represent the number of errors.
[a] $n = 27$.
[b] $n = 22$.

were equivalent, on average, to that of French-speaking children of 6 years of age ($SD = 1.28$), that is, 1;10 younger than the IA children (see Table 3).

In order to better understand the language test results of the IA children, a series of correlations were carried out between their language test scores and a number of age, health, and other related variables; these results are described now.

Correlations among language scores and health problems, speech–language services, and recruitment cohort

To further explore a possible link between the IA children's language test results and health issues, correlations were calculated between their language test scores and the presence of developmental or health problems, either in the past or at the time of testing. Accordingly, each IA child was assigned a score of 0 or 1 depending on whether their parents reported that they had *no* (0) or *some health problems* (1), either previously or at the time of testing. Results indicated that there were no significant correlations between the presence of reported developmental or health problems experienced by the IA children, either in the past or at the time of testing, and their performance on any of the language measures.

Additional correlations were carried out to ascertain whether there was a relationship between the prevalence of speech, language, academic, or behavioral

problems and the IA children's language results. Thirteen IA children (the "learning difficulty" subgroup) who were reported by teachers or parents to have either: (a) speech and language therapy or (b) academic or behavioral difficulties, including attention-deficit/hyperactivity disorders, were identified and assigned a code of "1." The remaining 15 children (the "typical learners" subgroup) were assigned a code of "0." Analyses revealed that there were no significant correlations between the IA children's learner status and their performance on any of the language tests.

Finally, to examine whether the low language test performance of the IA children could be attributed to differences between the IA children who had participated in Gauthier and Genesee's study (IA_C, continuing) and the newly recruited children (IA_N, new), each IA child was classified as "continuing" and assigned a code of "0" or "newly recruited" and assigned a code of "1." Correlations were then carried out between these subgroup codes and language test scores. Results for all these correlations are presented in Table 4. The only significant correlations that were found were between recruitment subgroup and age at testing ($r = -.53$, $p < .01$), length of exposure to French ($r = -.39$, $p = .04$), and performance on the word definition subtest ($r = -.44$, $p = .02$). These correlations indicated that the new IA children were younger at the time of this testing ($M = 7;6$ years) and, thus, had had less exposure to French ($M = 6;6$), and scored significantly lower on the word definition subtest ($M = 18.1$) than the continuing IA children from Gauthier and Genesee's study ($M_{age} = 8;1$, $M_{exposure} = 6;11$, $M_{word\ definition} = 23.8$). Because significant differences between the IA cohorts were restricted to the word definition subtest, cohort differences are not considered further.

Correlations between language scores and age-related variables

Correlations were calculated between age at testing, age at adoption, and length of exposure to French and the IA children's scores on the language tests (recalling sentence and the word association subtests of the Clinical Evaluation of Language Fundamentals—Revised, ECOSSE, EOWPVT, WIAT, EVIP, word definition subtest of the Wechsler Intelligence Scale for Children) in order to examine if differences in these age-related variables were related to differences among the IA children on the language tests. The age at which adopted children produced their first words in French postadoption was also correlated with their language scores because Gauthier and Genesee (2011) found it to be a significant, and in fact the most significant, predictor of adoptees' expressive language abilities at 5 years of age. Our results showed that age at adoption was significantly and negatively correlated with exposure to French ($r = -.59$, $p < .01$), as one would expect (see Table 4). Thus, age at adoption and amount of exposure to French are confounded, but there is no way to disentangle the effects of these variables using the present data. Unlike in Gauthier and Genesee, amount of exposure to French, and therefore age at adoption, was not significantly correlated with performance on the EVIP or the EOWPVT, although both correlated significantly with performance on the word associations ($r = .39$, $p = .04$) and word definitions ($r = .54$, $p < .01$) subtests; these two subtests test higher order vocabulary skills and were

Table 4. *Correlations between outcome variables and language measures*

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Subgroup of IA children	—	.03	−.07	−.39*	−.53**	.01	−.10	−.23	−.19	−.21	−.09	−.17	−.30	−.04	−.44*
2. Presence of difficulties		—	−.17	.20	.14	−.14	.04	.20	−.15	−.05	−.28	−.26	.07	−.05	.01
3. Age at adoption			—	−.59**	−.09	.81**	−.10	.18	.10	−.15	−.19	−.01	−.16	−.29	−.32
4. Exposure to French				—	.86**	−.34	.10	.03	.25	.39*	−.21	.19	.21	.32	.54**
5. Age at testing					—	.05	.19	.16	.36	.39*	−.14	.22	.36	.21	.46*
6. Age of first words in French						—	.16	−.11	.14	−.17	−.005	.08	.15	−.17	−.23
7. No. of health problems (past)							—	.41*	.13	.08	.06	.08	−.07	.05	−.20
8. No. of health problems (present)								—	.21	−.07	−.19	−.23	.07	.05	−.09
9. Recalling sentences (CELF-R)									—	.45*	.55**	.52*	.44*	.53*	.40*
10. Word associations (CELF-R)										—	.34	.49*	.43*	.52**	.42*
11. ECOSSE											—	−.54**	−.58**	−.41*	−.36
12. EOWPVT												—	.30	.23	.61**
13. EVIP													—	.39*	.35
14. WIAT-II														—	.52**
15. Word definition (WISC-IV)															—

Note: Subgroup of IA children, continuing internationally adopted children (IA_C) versus newly recruited IA children (IA_N); Presence of difficulties, low academic performers vs. typical academic performers; CELF-R, Clinical Evaluation of Language Fundamentals—Revised; ECOSSE, Épreuve de Compréhension Syntaxico-Sémantique; EOWPVT, Expressive One-Word Picture Vocabulary Test—Third Edition; EVIP, Échelle de Vocabulaire en Images Peabody; WIAT-II, Wechsler Individual Achievement Test—Second Edition; WISC-IV, Wechsler Individual Scale for Children—Fourth Edition.
*$p < .05$. **$p < .01$.

not administered by Gauthier and Genesee. Like Gauthier and Genesee, we did not find significant correlations between age at adoption or amount of exposure to French and performance on the ECOSSE, a test of grammatical competence. In contrast to Gauthier and Genesee's results, age at which adoptees produced their first words in French postadoption was not significantly correlated with any of the language scores.

In brief, the only significant correlations that emerged between amount of exposure to French and age of adoption involved tests of higher order knowledge of words. Moreover, and unlike Gauthier and Genesee (2011), we found that delay in production of first words in French was not correlated with any of our language measures. Thus, with the exception of word associations and word definitions, the age-related variables, all of which pertain to the early language experiences of the IA children, did not correlate with the language outcome measures in the present study.

Relationships between language scores and verbal memory

In order to explore the relationship between the IA children's performance on the language tests and the recalling sentences subtest, a measure that has been found by others to be sensitive to age of acquisition and learner status and is thought to assess aspects of verbal memory (both short and long term), additional correlations were calculated between scores on the recalling sentences subtest and the IA children's language test scores. The results revealed that performance on the recalling sentences subtest was significantly correlated with all of the other language measures: correlations varied from .40 ($p < .05$) for word definitions to .55 ($p < .01$) for receptive grammar. To investigate the relative importance of performance on the recalling sentences subtest in comparison to age at adoption, length of exposure to French, and age at which first words were produced in French, we ran a series of stepwise linear regression analyses using scores on each of these measures as predictor variables; scores on each of the languages tests were entered as outcomes variables. The results of these analyses revealed that the IA children's performance on the recalling sentences subtest was a significant predictor of their scores on the following tests: ECOSSE ($R^2 = .35, \beta = -0.60, p = .01$), EOWPVT ($R^2 = .29, \beta = 0.54, p = .03$), WIAT ($R^2 = .40, \beta = 0.43, p = .04$), EVIP ($R^2 = .23, \beta = 0.48, p = .03$), and word associations ($R^2 = .24, \beta = 0.43, p = .05$). The beta value is negative for the ECOSSE because scores represent the number of errors. Age at adoption was also a significant predictor of performances on the WIAT ($R^2 = .21, \beta = -0.49, p = .02$), while length of exposure to French was a significant predictor only of performance on the word definition subtest ($R^2 = .24, \beta = 0.49, p = .03$). These results suggest that reading comprehension and knowledge of word definitions are best predicted by IA children's age at adoption and length of exposure to French, respectively, whereas IA children's performance on the recalling sentences subtest is an important predictor of their performance on all the measures of language abilities, except the word definition subtest.

The IA and CTL children's results on the language tests on which they differed (namely, EOWPVT, ECOSSE, and word definitions) were compared again using a multivariate analysis of covariance ($\alpha = 0.05$), using recalling sentences subtest scores as a covariate to see whether significant differences between the IA and CTL children's performance on these would persist once the effects of performance on the test of sentence recall were partialled out. The results showed that there were no longer any significant group differences on any of these tests: EOWPVT, F (1, 47) = 0.09, $p = .76$; ECOSSE, F (1, 52) = 0.20, $p = .66$; and word definitions, F (1, 52) = 0.11, $p = .75$; once performance on the recalling sentences subtest was partialed out.

Taken together, these results suggest that the differences between the IA and CTL children on the language tests may be due to differences in their performance on the recalling sentences test which, as suggested by Gauthier and Genesee (in press), may in turn be related to their delayed onset of exposure to French, a point we return to in the Discussion section.

DISCUSSION

The purpose of the present study was to conduct a 3-year follow-up of IA children from China who were being raised by French-speaking parents and who had been tested initially by Gauthier and Genesee when they were 4;2 on average. More specifically, the study sought to examine if the IA children in Gauthier and Genesee's (2011) study would close the gap in language performance they exhibited relative to nonadopted comparison children with extended exposure to French into the early school years. Only 12 of Gauthier and Genesee's original sample agreed to participate in the present study, owing largely to time constraints, and thus an additional 15 IA children were recruited to increase the total sample size to 27. The IA children's results on a battery of language tests were compared to those of nonadopted monolingual French-speaking children who were carefully matched for age, gender, and SES. The inclusion of new IA children made it possible to examine the generalizability of Gauthier and Genesee's results to a new cohort of IA children. The IA children's socioemotional, health, and nonverbal cognitive development were also assessed and compared to that of the nonadopted children.

In terms of general health, the adoptive parents reported more health problems than did the parents of the nonadopted children. However, most of these problems were accounted for by vision difficulties, which were in no case severe. Moreover, although adoptive parents reported more socioemotional and behavioral difficulties than control parents on the Developmental Questionnaire, no significant differences were found between IA and CTL children on the CBCL and CBCL-TRF, two standardized measures. Results from the Wechsler Non-Verbal IQ test indicated that the IA children did not differ from the controls with respect to nonverbal cognitive abilities and, in particular, visuospatial memory, a result we return to later. Overall, these results indicate that the general physical, socioemotional and cognitive development of the IA children was comparable to

that of matched nonadopted peers, indicating both their resilience in overcoming possible preadoptive deprivations and remarkable progress in general development since adoption. Because the conditions of institutionalization of the IA children are unknown, these results could also be interpreted to indicate that the adoptees' preadoption environments were not severely impoverished or that any preadoption impoverishment they had experienced had limited and short-term effects. This latter interpretation corroborates the findings of previous studies that have found that, in general, Chinese adoptees do not suffer from health or general developmental difficulties postadoption (e.g., Johnson et al., 1998; Pomerleau et al., 2005).

In terms of language outcomes, the primary focus of the present study, the IA children scored within the normal range on all language tests, except the recalling sentences subtest. These findings replicate those reported by Gauthier and Genesee (2011). These results reinforce the point that any preadoption deprivation that the IA children might have experienced was either relatively minor or had limited and short-term effects on the children's language development and corroborate Gauthier and Genesee's and other's findings that, with sufficient exposure to their adoption language, IA children can develop skills in their new language that are within the normal range for typically developing average children (e.g., Cohen et al, 2008; Scott et al., 2008).

At the same time, our results indicate that there were important differences in language outcomes between the IA and CTL children. The IA children had significantly lower scores on measures of expressive vocabulary, receptive grammar, word definitions, and sentence recall. Moreover, the IA children exhibited larger within-sample variation than the control group and, in fact, a substantial percentage of IA children scored more than 1 *SD* below the mean of the control group on all language measures, although some of the IA children performed at the same level or higher than the CTL children. Large variation in language test performance was reported by Gauthier and Genesee (2011) and has been reported by others (e.g., Dalen & Ryvgold, 2006; Roberts et al., 2005). In short, a significant number of IA children had difficulty attaining parity with nonadopted peers from similar SES backgrounds, corroborating Gauthier and Genesee's findings, and those of Cohen and her colleagues, indicating that their results are not specific to the cohort of children they tested.

As just noted, adverse effects associated with preadoption adversity or impoverishment seem unlikely explanations of the differences in language outcomes for the two groups, as do effects related to differences in cognitive, including academic, and socioemotional development because, in fact, the IA children performed as well as the CTL children in these domains. In corroboration with Gauthier and Genesee's (2011) argument, insufficient exposure to French also seems an unlikely explanation because this is the third evaluation of these children that has provided evidence of lags in the IA children and, moreover, there is no evidence from the present study that the gap between the IA and CTL children was diminishing even after 3 years of additional exposure to French. Delayed onset of exposure to French, that is, early age effects, remains a viable possibility. In this regard, the results of the IA children on the recalling sentences subtest are of particular interest.

Performance on tests of sentence recall has been shown to be sensitive to age of acquisition effects in both L1 and L2 learners (French & O'Brien, 2008; Jessop, Suzuki & Tomita, 2007; Mayberry & Fisher, 1989; Rosselli et al., 2000; Vinther, 2002), and it is, in part, for this reason that we are hypothesizing that it is differences in age-related effects on performance on tests of sentence recall that might be mediating the differences in language outcomes between the IA and CTL children. Evidence in support of this possibility comes from the results of our regression analyses and multivariate analyses of covariance that indicated that scores on the recalling sentences subtest were significantly correlated with all other language test results and, in fact, no significant differences between the IA and CTL children on the language tests remained once performance on sentence recall was covaried out. If the failure of the IA children to close the language gap as a result of increased exposure to French is linked to their performance on the recalling sentences subtest, one might also expect to find little or no improvement in their performance on this test over time. In order to examine this possibility, we compared the recalling sentences results of the continuing IA children from Gauthier and Genesee's (2011) study ($n = 12$) with their results in the present study and found no significant improvement over time (Gauthier and Genesee: $M = 31.2$; present results: $M = 38.8$; $t = 1.49$, $p = .20$). In fact, 4 of the 12 continuing IA children had lower scores on the recalling sentences subtest at 7;0–8;0 than at 4;0–5;0, and 2 had scores that did not improve at all from one assessment to another. In contrast, the CTL children in the present study scored significantly higher than the CTL children ($n = 23$) in Gauthier and Genesee: mean of present controls = 51.1; mean of Gauthier and Genesee controls = 43.3; t (22) = 2.29, $p < .05$.

The IA children's results on the recalling sentences subtest contrast with their results on two other language tests that were administered in the present study as well as by Gauthier and Genesee (2011). Specifically, there were significant improvements over time on both the EOWPVT, $n = 6$ (Gauthier and Genesee: $M = 53.3$; present results: $M = 81.7$; $t = 7.27$, $p = .001$) and the EVIP, $n = 12$ (Gauthier and Genesee: $M = 78.5$; present results: $M = 114.0$; $t = 5.93$, $p < .001$). The IA children's longitudinal results on the recalling sentences subtest also contrast with results of age-related improvements reported by Gallimore and Tharp (1981) in an 8-year longitudinal study of sentence recall performance in typically developing monolingual English-speaking children. They found statistically significant age-related improvements in sentence recall performance between 5 and 8 years of age that were reliable across several cohorts of children who participated in their study. In sum, the IA children's performance on the recalling sentences subtest has been consistently low. Unlike their performance on the other language measures, it did not demonstrate significant improvement from the previous to the present assessment, contrary to what one would expect (Gallimore & Tharp, 1981). In addition, differences in language outcomes between the IA and CTL children were not found once differences on the recalling sentences subtest were taken into account statistically.

The question arises: What aspect of performance on the recalling sentences subtest might account for these results? Although, as mentioned earlier, tests

of sentence recall are often used in clinical assessments of children's language abilities (Archibald, Joanisse & Shepherd, 2008) and, indeed, performance on tests of sentence recall has been found to be a consistent and significant correlate of language impairments and dyslexia (Alloway & Gathercole, 2005), there is uncertainty as to what specific skills these tests involve. Several studies suggest that they assess children's knowledge of language in long-term memory, including grammar and vocabulary (Alloway & Gathercole, 2005; Gallimore & Tharp, 1981). There is also evidence that they involve auditory or phonological short-term memory, especially when tests are used that include sentences of increasing length, as was the case in the present study (Alloway & Gathercole, 2005). Thus, performance on tests of sentence recall seem to entail the integration of syntactic and lexical knowledge of language from long-term memory along with phonological information that has to be retained in short-term memory (Alloway & Gathercole, 2005; Baddeley, 2000; Gallimore & Tharp, 1981; Potter & Lombardi, 1998). That the IA children in the present study performed within age-expected levels on all of the other language tests, including tests related to vocabulary and grammar, suggests that the nature of their difficulties on the recalling sentences subtest is more related to short-term memory than to knowledge of language stored in long-term memory. The results from the present study also suggest that any putative gap between the IA and CTL children with respect to short-term memory is specific to verbal memory: there were no significant differences between the IA and CTL children on tests in the present battery that tapped other components of memory, namely, the recognition and spatial span subtests of the Wechsler Non-Verbal IQ test. Thus, unlike children with low working memory abilities who have difficulties with both verbal and visuospatial short-term memory (Alloway et al., 2009), the IA children's difficulties were limited to the verbal domain. Of course, these possibilities are largely speculative and can be elucidated only with more intensive investigations of the underlying components of sentence recall performance. In fact, we have such a study under way at the moment.

In conclusion, we found gaps in the language outcomes of IA children in the present study when compared to same-age CTL children that replicate those reported by Gauthier and Genesee (2011) from two previous assessments. These gaps, although statistically significant are subtle because they are revealed only when IA children are compared to carefully matched comparison groups. In fact, the IA children scored in the typical range for their age group on all of the language tests, except for the recalling sentences subtest, indicating that the gaps in language outcomes revealed in the present study and in previous work by Genesee and Gauthier and Cohen and colleagues are nonclinical in nature. Detailed analyses of the present results suggest the intriguing possibility that delayed age of onset of acquisition of French may limit IA children's verbal short-term memory, which in turn may affect their language-learning outcomes. It could also be that disruption of the children's acquisition of the birth language or a combination of disruption and delay is at issue. Whatever the contextual factor(s) at play, there is also the question of what psycholinguistic factor or factors account for these effects. We have proposed that verbal short-term memory is a possible mitigating factor. Clearly much additional

research is needed to substantiate these possibilities or, alternatively, to reveal other explanations.

Limitations

Although the sample size in the present study was large enough to yield statistically significant differences and patterns, including more children would yield even more reliable results. Only IA children from China were included because of their generally good overall development postadoption, making differences in their language development of particular interest from the point of view of age of acquisition effects. Nevertheless, replication of the present study with IA children from other countries and language backgrounds would serve to establish the generalizability of the present results.

APPENDIX A

Demographic data of the IA and CTL children

	IA	CTL
Age (years;months), *M* (*SD*)	7;10 (6.00 months)[a]	7;11 (6.95 months)[a]
Age at adoption (years;months), *M* (*SD*)	1;1 (3.82 months)[a]	
Length of exposure to French (years;months), *M* (*SD*)	6;9 (7.39 months)[a]	7;11 (6.95 months)[a]
Mother's age (years), *M* (*SD*)	46.70 (5.40)[b]	40.57 (5.08)[c]
Father's age (years), *M* (*SD*)	48.63 (5.66)[d]	42.57 (6.26)[c]
Mother's level of education, % (*n*)		
High school	7.4 (2)	11.1 (3)
College	33.3 (9)	22.2 (6)
University	59.3 (16)	66.7 (18)
Father's level of education, % (*n*)		
High school	23.1 (6)	14.8 (4)
College	34.6 (9)	29.6 (8)
University	42.3 (11)	55.5 (15)
Family income per year, % (*n*)		
$30,000–59,000	7.4 (2)	3.7 (1)
$60,000–89,000	18.5 (5)	22.2 (6)
$90,000–119,000	33.3 (9)	25.9 (7)
$120,000–150,000	18.5 (5)	14.8 (4)
$150,000 and more	22.2 (6)	33.3 (9)

Note: The level of education of one of the fathers was not disclosed for personal reasons. IA, Internationally adopted; CTL, control.
[a] $n = 27$.
[b] $n = 20$.
[c] $n = 14$.
[d] $n = 19$.

APPENDIX B

Health and developmental problems

	Past		Present	
	IA	CTL	IA	CTL
Health and/or Developmental Problems	% (*n*)	% (*n*)	% (*n*)	% (*n*)
Gross motor delay	7.4 (2)	0 (0)	0 (0)	0 (0)
Mild developmental delay	3.7 (1)	0 (0)	0 (0)	0 (0)
ADHD	3.7 (1)	0 (0)	11.1 (3)	0 (0)
Infectious or parasitic disease	3.7 (1)	0 (0)	0 (0)	0 (0)
Ear infections	22.2 (6)	81.5 (22)	0 (0)	7.4 (2)
1 and 3 otitis	22.2 (6)	29.6 (8)	0 (0)	7.4 (2)
3 and 5 otitis	0 (0)	18.5 (5)	0 (0)	0 (0)
5 and more otitis	0 (0)	14.8 (4)	0 (0)	0 (0)
Frequency unspecified	0 (0)	18.5 (5)	0 (0)	0 (0)
Hearing difficulties	0 (0)	0 (0)	0 (0)	0 (0)
Respiratory difficulties	3.7 (1)	22.2 (6)	14.8 (4)	22.2 (6)
Vision impairments	7.4 (2)	3.7 (1)	29.6 (8)	3.7 (1)
Emotional and/or attachment difficulties	18.5 (5)	3.7 (1)	14.8 (4)	0 (0)
Behavioral difficulties	14.8 (4)	0 (0)	7.4 (2)	0 (0)
Weight or height below 10th percentile	14.8 (4)	11.1 (3)	0 (0)	0 (0)
Total instances of reported problems	27	33	21	9

Note: No cases of physical disability, neurological problems, or fetal alcohol syndrome were reported. IA, Internationally adopted; CTL, control; ADHD, attention-deficit/hyperactivity disorder.

ACKNOWLEDGMENTS

This research was supported by a grant from the Social Sciences and Humanities Research Council of Canada (to F.G.). We thank the adoption agencies Formons une famille Inc. and Parents sans frontières Inc. for their help with the recruitment of IA children and the elementary schools (Collège Français, École Les Trois Saisons, Le Pensionnat des Sacrés Coeurs, L'Académie Louis Pasteur) for their help with the recruitment of nonadopted monolingual French-speaking children. We express our thanks to our participants and their parents.

REFERENCES

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

Achenbach, T. M. (1991a). *Manual for the Child Behavior Checklist/4–18 and 1991 Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M. (1991b). *Manual for the Teacher's Report Form and 1991 Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Alloway, T., Gathercole, S., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development, 80*, 606–621.

Alloway, T. P., & Gathercole, S. E. (2005). The role of sentence recall in reading and language skills of children with learning difficulties. *Learning and Individual Differences, 15*, 271–282.

Andresen, I. L. (1992). Behavioral and school adjustment of 12–13-year old internationally adopted children in Norway: A research note. *Journal of Child Psychology and Psychiatry, 33*, 427–439.

Archibald, L. M. D., Joanisse, M. F., & Shepherd, M. (2008). Associations between key language-related measures in typically developing school-age children. *Journal of Psychology, 216*, 161–171.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*, 417–422.

Baddeley, A. D. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*, 189–208.

Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language, 44*, 235–249.

Bishop, D. V. (1983). *Test for Reception of Grammar*. Manchester, UK: Author.

Brownell, R. (2000). *Manual for expressive one-word picture vocabulary test*. Novato, CA: Academic Therapy Publications.

Clark, E. A., & Hanisee, J. (1982). Intellectual and adaptive performance of Asian children in adoptive American settings. *Developmental Psychology, 18*, 595–599.

Cohen, N. J., Lojkasek, M., Zadeh, Z. Y., Pugliese, M., & Kiefer, H. (2008). Children adopted from China: A prospective study of their growth and development. *Journal of Child Psychology and Psychiatry, 49*, 458–468.

Conti-Ramsden, G. M., Botting, N. F., & Faragher, B. (2001). Psycholinguistic markers of specific language impairments (SLI). *Journal of Child Psychology and Psychiatry, 42*, 741–748.

Croft, C., Beckett, C., Rutter, M., Castle, J., Colvert, E., Groothues, C., et al. (2007). Early adolescent outcomes of institutionally-deprived and non-deprived adoptees: Language as a protective factor and a vulnerable outcome. *Journal of Child Psychology and Psychiatry, 48*, 31–44.

Dalen, M., & Rygvold, A.-L. (2006). Educational achievement in adopted children from China. *Adoption Quarterly, 9*, 45–58.

DeKeyser, R. M. (2000). The robustness of critical period effects on second language acquisition. *Studies in Second Language Acquisition, 22*, 499–533.

Dunn, L. M., Theriault-Whalen, C. M., & Dunn, L. M. (1993). *Manual for échelle de vocabulaire en images Peabody*. Toronto: Psycan Corporation.

French, L. M., & O'Brien, I. (2008). Phonological memory and children's second language grammar learning. *Applied Psycholinguistics, 29*, 463–487.

Gallimore, R., & Tharp, R. G. (1981). The interpretation of elicited sentence imitation in a standardized context. *Language Learning, 31*, 369–392.

Gathercole, S. (n.d.). *The structure and functioning of phonological short-term memory*. Retrieved April 10, 2010, from http://www.acfos.org/publication/ourarticles/pdf/acfos2/gathercole.pdf

Gauthier, K., & Genesee, F. (2011). Language development in internationally-adopted children: A special case of early second language learning. *Child Development, 82*, 1–15.

Geren, J., Snedeker, J., & Ax, L. (2005). Starting over: A preliminary study of early lexical and syntactic development in internationally adopted preschoolers. *Seminars in Speech and Language, 26*, 44–53.

Glennen, S. (2002). Language development and delay in internationally adopted infants and toddlers. *Journal of Speech–Language Pathology, 11*, 333–339.

Glennen, S. (2009). Speech and language guidelines for children adopted from abroad at older ages. *Topics in Language Disorders, 29*, 50–64.

Glennen, S., & Bright, B. J. (2005). Five years later: Language in school-age internationally adopted children. *Seminars in Speech and Language, 26*, 86–101.

Glennen, S., & Masters, G. (2002). Typical and atypical language development in infants and toddlers adopted from Eastern Europe. *American Journal of Speech–Language Pathology, 11*, 417–433.

Groze, V., & Ileana, D. (1996). A follow-up study of adopted children from Romania. *Child and Adolescent Social Work Journal, 13*, 541–565.

Gunnar, M. R., Bruce, J., & Grotevant, H. D. (2000). International adoption of internationally reared children: Research and policy. *Development and Psychopathology, 12*, 677–693.

Hellerstedt, W., Madsen, N., Gunnar, M., Grotevant, H., Lee, R., & Johnson, D. (2008). The International Adoption Project: Population-based surveillance of Minnesota parents who adopted children internationally. *Maternal and Child Health Journal, 12*, 162–171.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*, 1368–1378.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review, 26*, 55–88.

Hoff, E., & Tian, C. (2005). Socioeconomic status and cultural influences on language. *Journal of Communication Disorders, 38*, 271–278.

Hôpital Ste-Justine. (2009). *L'otite.* Retrieved March 20, 2010, from http://www.chu-sainte-justine.org/Famille/page.aspx?ID_Menu=668&ID_Page=1595&ItemID=3a

Hwa-Froelich, D. A., & Matsuoh, H. (2008). Cross-cultural adaptation of internationally adopted Chinese children. *Communication Disorders Quarterly, 29*, 149–165.

Hyltenstam, K. (1992). Non-native features of near-native speakers. On the ultimate attainment of childhood L2 learners. In R. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 351–368). Amsterdam: Elsevier.

Hyltenstam, K., Bylund, E., Abrahamsson, N., & Park, H.-S. (2009), Dominant-language replacement: The case of international adoptees. *Bilingualism: Language and Cognition, 12*, 121–140.

Jessop, L., Suzuki, W., & Tomita, Y. (2007). Elicited imitation in second language acquisition research. *Canadian Modern Language Review, 64*, 215–220.

Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, 60–99.

Johnson, K., Banghan, H., & Liyao, W. (1998). Infant abandonment and adoption in China. *Population and Development Review, 24*, 469–510.

Kleinstein, R. N., Jones, L. A., Hullett, S., Kwon, S., Lee, R. J., Friedman, N. E., et al. (2003). Refractice error and ethnicity in children. *Archives of Ophthalmology, 121*, 1141–1147.

Krakow, R., Tao, S., & Roberts, J. (2005). Adoption age effects on English language acquisition: Infants and toddlers from China. *Seminars in Speech and Language, 26*, 33–43.

Lecocq, P. (1996). *Épreuve de compréhension syntaxico-sémantique*. Villeneuve d'Ascq, France: Presses Universitaires du Septentrion.

LeNormand, M. T., Parisse, C., & Cohen, H. (2008). Lexical diversity and productivity in French preschoolers: Developmental, gender, and sociocultural factors. *Clinical Linguistics and Phonetics, 22*, 47–58.

Mayberry, R. I., & Eichen, E. B. (1991). The long-lasting advantage of learning sign language in childhood: Another look at the critical period for language acquisition. *Journal of Memory and Language, 30*, 486–512.

Mayberry, R. I., & Fisher, S. D. (1989). Looking through phonological shape to lexical meaning: The bottleneck of non-native sign language processing. *Memory and Cognition, 17*, 740–754.

Meacham, A. N. (2006). Language learning and the internationally adopted child. *Early Childhood Education Journal, 34*, 73–79.

Paradis, J., Genesee, F., & Crago, M. (in press). *Dual language development and disorders*. Baltimore, MD: Brookes.

Penfield, W., & Robert, L. (1959). *Speech and brain mechanisms*. Princeton, NJ: Princeton University Press.

Pollock, K. (2005). Early language growth in children adopted from China: Preliminary normative data. *Seminars in Speech and Language, 26*, 22–32.

Pomerleau, A., Malcuit, G., Chicoine, J. F., Séguin, R., Belhumeur, C., Germain, P., et al. (2005). Health status, cognitive, and motor development of young children adopted from China, East Asia, and Russia across the first 6 months after adoption. *International Journal of Behavioral Development, 29*, 445–457.

Potter, M. C., & Lombardi, L. (1998). Syntactic priming in immediate recall of sentence. *Journal of Memory and Language, 38*, 265–282.

Roberts, J., & Krakow, R. (2003). Language outcomes for preschool children adopted from China as infants and toddlers. *Journal of Multilingual Communication Disorders, 1*, 177–183.

Roberts, J., Krakow, R., & Pollock, K. (2003). Language outcomes for preschool children adopted from China as infants and toddlers. *Journal of Multilingual Communication Disorders, 1*, 177–183.

Roberts, J., Pollock, K., & Krakow, R. (2005). Continued catch-up and language delay in children adopted from China. *Seminars in Speech and Language, 26*, 76–85.

Roberts, J., Pollock, K., Krakow, R., Price, J., Fulmer, K., & Wang, P. (2005). Language development in preschool-age children adopted from China. *Journal of Speech, Language and Hearing Research, 48*, 93–107.

Rojewski, J. W., Shapiro, M. S., & Shapiro, M. (2000). Parental assessment of behavior in Chinese adoptees during early childhood. *Child Psychiatry and Human Development, 31*, 79–96.

Rosselli, M., Ardila, A., Araujo, K., Weekes, V. A., Caracciolo, V., Padilla, M., et al. (2000). Verbal fluency and repetition skills in healthy older Spanish–English bilinguals. *Applied Neuropsychology, 7*, 17–24.

Scott, K., Roberts, J., & Krakow, R. (2008). Oral and written language development of children adoption from China. *American Journal of Speech–Language Pathology, 17*, 150–160.

Semel, E., Wiig, E. S., & Secord, W. A. (1987). *Clinical evaluation of language Fundamentals—Revised*. San Antonio, TX: Psychological Corporation.

Snedeker, J., Geren, J., & Shafto, C. (2007). Starting over: International adoption as a natural experiment in language development. *Association for Psychological Science, 18*, 78–87.

Stokes, S., Wong, A., Fletcher, P., & Leonard, L. (2006). Nonword repetition and sentence repetition as clinical markers of specific language impairment: The case of Cantonese. *Journal of Speech, Language, and Hearing Research, 49*, 219–236.

Tan, T. X., & Marfo, K. (2006). Parental ratings of behavioral adjustment in two samples of adopted Chinese girls: Age-related versus socio-emotional correlates and predictors. *Applied Developmental Psychology, 27*, 14–30.

Tan, T. X., & Yang, Y. (2005). Language development of Chinese adoptees 18–35 months old. *Early Childhood Research Quarterly, 20*, 57–68.

van IJzendoorn, M., Juffer, F., & Poelhuis, C. (2005). Adoption and cognitive development: A meta-analytic comparison of adopted and non-adopted children's IQ and school performance. *Psychological Bulletin, 131*, 301–316.

Vinther, T. (2002). Elicited imitation: A brief overview. *International Journal of Applied Linguistics, 12*, 52–73.

Wechsler, D. (2003). *Échelle d'Intelligence de Wechsler pour Enfants, Quatrième Édition—Version pour Francophones du Canada*. Toronto: Harcourt Assessment.

Wechsler, D. (2005). *Test de Rendement Individuel de Wechsler*. Toronto: Harcourt Assessment.

Wechsler, D., & Naglieri, J. (2006). *Wechsler Non-Verbal Scale of Ability*. Toronto: Harcourt Assessment.