Bilingualism: Language and Cognition 17 (3), 2014, 556–571 © Cambridge University Press 2013 doi:10.1017/S1366728913000746

Lexical-semantic skills in bilingual children who are becoming English-dominant: A longitudinal study*

LI SHENG
The University of Texas at Austin

(Received: May 20, 2012; final revision received: November 3, 2013; accepted: November 3, 2013; first published online 13 December 2013)

Twenty-seven Mandarin–English bilingual children participated in picture identification and picture naming tasks at two time points, 16 months apart. The younger children (mean age = 4 years) showed greater gains over time than the older children (mean age = 6 years 10 months) in English lexical-semantic skills and neither group showed significant gains in Mandarin. At the individual level, a majority of the children showed increased accuracy for the English tasks, but only half of them did so for the Mandarin tasks. Analyses of error distribution indicated production of more advanced error types in the older children and in English, as well as different patterns of time-related changes in error types in the two languages. These findings illustrate how age and initial language proficiency are related to lexical growth among Mandarin-speaking bilingual children who are becoming English-dominant.

Keywords: picture naming, picture identification, Mandarin, English, error analysis, longitudinal method

The increased representation of bilingual populations across the world calls for in-depth studies of the various circumstances in which children are becoming bilingual and how contextual (e.g., community, school, family, linguistic typology) and individual variables impact the development of different linguistic skills (Barac & Bialystok, 2012; Goldenberg, Reese & Rezaei, 2011; Han, 2008; Marinova-Todd & Uchikoshi, 2011; Sheng, Lu & Kan, 2011). The goal of the current study is to enhance our understanding of lexical-semantic skill development in bilingual children's first language (L1) and second language (L2) when the children are growing up in an L2-dominant environment. In particular, a group of Mandarin–English bilingual children who participated in a previous cross-sectional study (Sheng et al., 2011) were followed longitudinally and tested on tasks of receptive and expressive vocabulary using measures of accuracy and error types. In the following sections, I review previous studies of bilingual lexical development and discuss factors that may lead to cross-study differences in outcomes. This is followed by a brief review of lexical retrieval errors and the current research questions.

* The author would like to thank the children and their families for participating in the study, Ying Lu and Brooke Lauper for collecting the data, and Dr. Carmen Silva-Corvalán and three anonymous reviewers for their comments on previous versions of this paper. This research was supported by a Grant Preparation Award from the University of Texas at Austin.

Bilingual lexical development

Many studies of bilingual lexical development in the U.S. were conducted with children who speak Spanish as their home language. Earlier works with Spanish–English bilingual toddlers highlighted the need to test bilingual children in both languages (Pearson, Fernández & Oller, 1993, 1995). More recent studies have focused on bilingual children who were of preschool and school age (Cobo-Lewis, Pearson, Eilers & Umbel, 2002; Hammer, Lawrence & Miccio, 2008; Rodríguez, Díaz, Duran & Espinosa, 1995; Winsler, Díaz, Espinosa & Rodríguez, 1999). Spanish-speaking children growing up in the U.S. show swift growth in English vocabulary alongside significant albeit smaller gains in Spanish vocabulary. These patterns were observed for children from both low-income families (Hammer et al., 2008; Rodríguez et al., 1995; Winsler et al., 1999) and middle socioeconomic background (Cobo-Lewis et al., 2002). In two studies, Rodríguez et al. (1995) and Winsler et al. (1999) compared vocabulary growth between Spanish-speaking children who went to bilingual preschools and those who did not attend preschools. Rodríguez et al. (1995) measured the children's vocabulary twice over one year. Winsler et al. (1999) followed the children in Rodríguez et al. for another year and recruited a new cohort of children and followed them for one year for replication purpose. Children in both studies received high support of Spanish use outside of home and school in the communities they lived. Rodríguez et al. and Winsler et al. found no erosion of L1 due to English exposure and argued

Address for correspondence:
Communication Sciences and Disorders, University of Texas at Austin, 2504-A Whitis Ave., Austin, TX 78712, USA.
li.sheng@utexas.edu

that these children were experiencing a form of additive bilingualism, in the sense that the introduction of English did not stall the development of the L1, at least in terms of performance on verbal tasks during the preschool period.

In contrast, other studies reported signs of suppressed L1 development among Spanish-speaking children. Hammer et al. (2008) compared two groups of children attending monolingual Head Start program. One group had dual language exposure at home prior to Head Start while the other group was exposed to Spanish only at home. During the two years of Head Start, the gap between the two groups increased for Spanish vocabulary but narrowed for English vocabulary. Thus, attending monolingual preschool program had differential effects on bilingual children's L1 versus L2 vocabulary growth. In a similar vein, Cobo-Lewis et al. (2002) found that although home language practice (whether the child was exposed to Spanish only or to both Spanish and English in the home environment prior to kindergarten entry) did not have a lasting effect on the children's English vocabulary by fifth grade, it had a sustained effect on Spanish vocabulary such that children from Spanish-only households outperformed those from dual language households on Spanish vocabulary tasks throughout the kindergarten to fifth grade years.

Studies with bilingual children who speak a minority language other than Spanish are emerging. Kan and Kohnert (2005) conducted a cross-sectional study of Hmong–English preschoolers who were attending a bilingual preschool for low-income families two or three days a week. The group of older children had a mean age of five years and had attended the preschool for 16 months; the younger children had a mean age of three years 11 months and had attended the preschool for nine months. Children were assessed using examiner-designed expressive and receptive vocabulary tasks. The older group outperformed the younger group on the English tasks but the two groups did not differ on the Hmong tasks. Differences between picture identification (receptive vocabulary) and picture naming (expressive vocabulary) were greater in Hmong than in English. These findings indicated rapid expansion of English vocabulary alongside relative stabilization of L1 vocabulary. Nevertheless, the children's composite vocabulary was higher than their vocabulary in English or Hmong alone, suggesting that these children still knew many Hmong words that they did not know in English.

A more recent cross-sectional study of another minority Asian language revealed both similar and different findings. Sheng et al. (2011) tested two groups of Mandarin–English children: a group of younger children with a mean age of 4;5 (years; months) and a group of older children with a mean age of 7;2. Both groups were exposed to Mandarin from birth and began systematic English exposure at around two years of age. Unlike the children in many of the previous studies (Cobo-Lewis et al., 2002; Kan & Kohnert, 2005; Rodríguez et al., 1995; Winsler et al., 1999), who had regular access to L1 in preschool or school settings, the Mandarin–English children received little if any systematic L1 support in the school setting. The only available form of Mandarin instructional support was two hours of Sunday Chinese lessons that aimed to teach children to read and write Chinese characters. Children participated in examiner-designed picture identification and picture naming tasks in both languages. Similarly to Kan and Kohnert (2005), the older children scored significantly higher than the younger children on the English tasks, but not the Mandarin tasks. Also similarly to Kan and Kohnert (2005), the performance gap between the naming and identification tasks was larger for Mandarin than for English. Unlike in Kan and Kohnert (2005), composite vocabulary was equal to English vocabulary and both were larger than Mandarin vocabulary. This divergent finding may be explained by the age difference in the study samples and the lack of systematic L1 support in schools for the Mandarin-speaking children. Sheng et al.'s (2011) findings provided evidence of rapid growth in English vocabulary and simultaneous early stagnation in L1 vocabulary in immigrant children growing up in an English-dominant environment. These patterns occurred against the backdrop of considerable L1 input in the home setting. On average, parents reported 43% of Mandarin input for the younger group and 41% of Mandarin input for the older group.

So far, studies on bilingual lexical development have included speakers of Spanish as well as minority Asian languages such as Hmong and Mandarin. However, differences in study design, sampling procedures, and measurements make direct comparisons across these studies difficult. A recent study by Uchikoshi (published online September 11, 2012) circumvents these problems and for the first time enables such a comparison. Uchikoshi conducted a three-year longitudinal study on vocabulary development in bilingual children whose L1 was, respectively, Spanish and Cantonese. Both groups of children were from working-class families and approximately half of the children attended transitional bilingual programs. Two differences were noted in the vocabulary growth patterns of these two groups. First, the Cantonese-speaking children had higher English expressive vocabulary scores than the Spanish-speaking children and this difference remained during the three-year span from kindergarten to second grade. Second, whereas the two groups showed comparable L1 receptive vocabulary scores at the beginning of kindergarten, the Spanish group showed steeper growth in L1 vocabulary than the Cantonese group.

Uchikoshi (published online September 11, 2012) ruled out access to L1 use opportunities as a potential reason for these differences because both groups of children lived in communities with high percentages of L1 speakers. Instead, age of first English exposure, parental expectations, and home language use were presented as potential contributors and areas of future investigation. In addition, as different groups of researchers have argued, bilingual language development is affected by a multitude of contextual and individual level variables which include, but are not limited to, overall language exposure and use, community language practice and home language practice, degree of linguistic overlap between L1 and L2, access to L1 in educational settings, and the child's own developmental level and cognitive ability (Bohman, Bedore, Peña, Mendez-Perez & Gillam, 2011; Goldenberg et al., 2011; Marinova-Todd & Uchikoshi; 2011; Sheng et al., 2011).

While these factors are frequently used to elucidate growth patterns at the group level, few studies have examined growth trajectory at the individual level. One notable exception is Kohnert (2002), who tested a group of 28 early sequential Spanish–English bilingual children on a picture naming task at two time points, 13 months apart. The children ranged in age from 5;9 to 13;10 at the initial testing ($M = 9.1$ years). The target words in the naming task involved familiar vocabulary because the researcher was more interested in the processing of known words than the acquisition of new words. The naming task consisted of two single-language conditions and one mixed-language condition. In the English-only condition, children named a block of 25 pictures presented on a computer screen with a simultaneous auditory cue (the English word *say* to indicate English naming trials). In the Spanish-only condition, children named a block of 25 pictures presented with the Spanish word *diga* "say" to indicate Spanish naming trials. In the mixed-language condition, children named a block of 50 pictures with auditory cues (*say* or *diga*) alternating languages on every third trial. Both naming accuracy and speed were examined. At the group level, accuracy and speed of naming in Spanish remained largely unchanged at the two time points. At the same time, gains in English naming accuracy and speed were greater and particularly significant in the high competition mixed-language condition. These results reinforced the author's previous cross-sectional findings of a shift toward greater strength in L2 lexical processing with age and increased language experience (Kohnert, Bates & Hernandez, 1999).

Further analyses of individual children's performance in Kohnert (2002) indicated a complex non-monotonic pattern of L1/L2 change over time. Specifically, no child demonstrated gains on all four dependent variables; most children experienced some gains, some losses, and some stability in lexical retrieval performance over time. Individual variability notwithstanding, the advantage of English over Spanish in growth patterns was apparent. For instance, 20 of the 28 children showed increased accuracy in English naming in the single-language condition, whereas only 10 of the 28 children showed improvement in Spanish on the accuracy measure.

In summary, studies of bilingual lexical development revealed steeper growth in English than L1 vocabulary in children who spoke Spanish, Cantonese, Hmong, and Mandarin. Growth of the L1 varied between different L1 groups and is determined by multiple factors. Moreover, bilingual lexical development was characterized by great inter- and intra-individual variability. These previous studies have mostly utilized measures of lexical retrieval accuracy and occasionally, speed. An area largely neglected is the errors bilingual children make in lexical tasks. This is the focus of the next section.

Lexical retrieval errors

Studies of monolingual children's picture naming reported several common error types such as semantic substitutions, phonological substitutions, indeterminate (or "don't know") errors, and unrelated errors (German & Newman, 2004; McGregor, 1997; Sheng & McGregor, 2010). Semantic and "don't know" errors are the most frequent error types whereas phonological errors are relatively rare.

Naming errors in clinical populations help to inform us about the developmental maturity of different error types. Children with specific language impairment and/or word finding deficits make more phonological and unrelated errors and fewer semantic errors in comparison to typical peers (McGregor, 1997; Sheng & McGregor, 2010). The error profiles – elevated level of phonological and unrelated errors, combined with lower level of semantic errors – demonstrated by these children were very similar to those of vocabulary-matched peers who were on average 19 months younger (Sheng & McGregor, 2010). It remains to be seen if bilingual children may demonstrate different error profiles in their two languages that align with different developmental stages.

In a study of French–English bilingual children (ages 7;0 to 10;1), Yan and Nicoladis (2009) examined lexical-semantic skills using standardized tests of receptive vocabulary and a researcher-designed picture naming task. They found that when unable to name pictures in the specified language, children frequently produced a more general word, a word from the same semantic domain, coined names, or the correct name from the other language. In studies of second language learners (Dörnyei & Kormos, 1998; Dörnyei & Scott, 1997; Greene, Bedore & Peña, 2012), researchers found that use of general-all-purpose words (e.g., *thing*, *stuff*, or *bird* in place of *duck*), semantic substitutions, circumlocutions (e.g., *ring*: "round like a circle, but you put it here [point to finger]"), word coinage (e.g., *branching* in place of *climbing*, *ocean ladder* in place of *bridge*), and code-switches in discourse contexts are common communicative strategies adopted by learners to solve lexical retrieval problems and to compensate for limited vocabulary. Finally, in a study of Spanish–English bilingual children's semantic depth, Sheng, Bedore, Peña and Fiestas (2013) found that children were much more likely to switch from Spanish to English than vice versa when producing word associations and a majority of the language-switched responses resulted in semantically related associations.

The current study

The purpose of the current study is to refine our knowledge of the developmental trajectory of Mandarin–English bilingual children's L1 and L2 lexicon as these children are becoming increasingly English-dominant. To achieve this goal, we utilized a longitudinal design and examined group patterns and individual profiles over time. Moreover, we provided detailed analyses of error types and tracked the changes in error profiles. The current study addressed the following questions:

1. What is the profile of L1 and L2 vocabulary development over time? Does the profile vary by age?
2. What patterns of change are seen at the individual level? Specifically, do children show increase, decrease, or no change in vocabulary task performance? Which of these three possible profiles is the most common for each of the two languages?
3. What factors, among age, cumulative language experience, language input and output, and general language proficiency as rated by parents, are related to vocabulary changes?
4. What errors do children produce when they are unable to retrieve a target? What are the effects of age, language, and time on error types?

On the basis of results from a previous cross-sectional study (Sheng et al., 2011), it was predicted that both younger and older bilingual children would show significant growth in English vocabulary and minimal growth in Mandarin vocabulary. At the individual level, given findings from Kohnert (2002), it was predicted that increase on all four measures of vocabulary (two languages × two tasks) would be relatively rare. Rather, children may show gains on certain measures, but loss or no change on other measures. However, it was predicted that gains would be more common for the English than the Mandarin measures.

With regard to factors related to vocabulary change, it was predicted that children who began with lower English exposure and proficiency would show the greatest improvement in English vocabulary within the 16 months period (Goldenberg et al., 2011). Given that greater exposure to and practice using the L1 were associated with greater L1 semantic knowledge (Bohman et al., 2011; Sheng et al., 2011), we predicted that Mandarin input and output measures would be correlated with changes in L1 vocabulary.

Finally, with regard to errors, it was predicted that more advanced errors (e.g., semantic errors) would be more common in older children, at time 2, and in children's stronger language; whereas less advanced errors (e.g., phonological, "don't know", and unrelated errors) would be more common in younger children, at time 1, and in the weaker language. Furthermore, language-switches were predicted to be common when responding in Mandarin, the home language, but relatively rare when responding in English, the school language (Sheng et al., 2013).

Method

Participants

Twenty-seven children (13 girls) participated in the same picture identification and picture naming tasks at two time points with an interval of 15.7 months[1] (SD = 1.6 months). To be included at time 1, children had to meet three inclusionary criteria: typically developing, spoke Mandarin and English on a daily basis, and had at least 20% input in the non-dominant language (see the parent interview passage below for details on how input percentages were calculated). Using the typical age of the onset of formal schooling – which is around five years of age in the United States – as the grouping criterion, the children were divided into a group of younger children ($n = 12$) and a group of older children ($n = 15$). At time 1, the younger children ranged between 3;1 and 5;0 (M = 4;0), and the older children ranged between 5;7 and 8;5 (M = 6;10). The three-to-eight-year-old age range was sampled because children make tremendous gains in oral language vocabulary during this period. Data from eight additional children (four in each group) who participated at time 1 were excluded because they could not be reached at time 2 due to family relocation or scheduling conflicts. Average years of maternal education for the younger and older groups was, respectively, 17.50 ($SD = 3.26$) and 17.27 ($SD = 2.76$), with no difference between the two means, $p > .50$. Average age of first English exposure was 26.50 months ($SD = 11.60$) for the younger and 23.67 months ($SD = 11.54$) for the older children, $p > .50$.

[1] We began re-testing the children after at least 12 months had elapsed because we expect early sequential bilingual children to make significant vocabulary gains after a one-year period (Kohnert, 2002). Every effort was made to re-test the children within a 12–18 months window, while accommodating the various schedules of participating families and the student testers. This was possible for all but one case.

Table 1. *Participant characteristics.*

Participant	Age (in months) at time 1	Age (in months) at time 2	Months from time 1 to time 2	Age (in months) at first English exposure
1	37	54	17	20
2	40	57	17	24
3	41	57	16	36
4	41	57	16	36
5	43	59	16	36
6	46	59	13	24
7*	48	62	14	34
8*	53	69	16	24
9	54	66	12	36
10*	57	73	16	12
11*	57	72	15	36
12*	60	76	16	0
13*	67	83	16	6
14*	71	90	19	0
15*	72	88	16	30
16*	72	86	14	36
17*	77	94	17	30
18	79	95	16	12
19	84	100	16	36
20*	84	98	14	36
21*	85	99	14	24
22	86	104	18	24
23*	87	104	17	24
24*	88	102	14	36
25*	90	107	17	11
26*	92	109	17	26
27*	101	115	14	24

* These children went to Sunday Chinese schools that offered approximately two hours of instruction per week on Chinese literacy skills.
Note: Children are arranged in order from youngest to oldest according to age at time 1. Participants #1–#12 belong to the younger age group and participants #13–#27 belong to the older age group. Age of first English exposure was estimated by parents retrospectively. All children acquired Mandarin as first language (from birth) in the home.

Information about individual participants' age and age of first English exposure is presented in Table 1.

The primary caregiver (a parent) of each child filled out a language and educational history questionnaire (based on Gutiérrez-Clellen & Kreiter, 2003; and Restrepo, 1998) at both time points. Each question on the questionnaire was printed in both English and Mandarin Chinese to facilitate the parents' understanding. After parents filled out the questionnaire, the examiner reviewed the questionnaire with the parents so that both parties had the opportunity to ask clarification questions if anything was unclear. Parents rated their children's oral language proficiency in the areas of grammar, sentence length, vocabulary, listening comprehension, and pronunciation using a five-point scale with one indicating low proficiency and five indicating high proficiency. Scores in these five domains were averaged for each child to derive an overall oral proficiency rating.[2]

[2] We ran correlation to see if parent ratings of the children's English proficiency correlated with children's performance on the English picture ID and picture naming tasks. At time 1, parent rating of their children's English proficiency correlated significantly with the children's English picture ID ($r = .82$, $p < .001$) and picture naming ($r = .78$, $p < .001$) performance. At time 2, parent rating of children's English proficiency correlated significantly with the children's English naming performance ($r = .38$, $p < .05$), but not with their picture ID performance ($r = .14$). The lack of correlation for the latter case was due to near-ceiling performance on the English picture ID task at time 2 (M = 95%, SD = 5%, range = 85–100%). These significant correlations suggest that the parents were able to provide valid estimates of their children's English proficiency. Parent ratings of children's Mandarin proficiency significantly correlated with children's Mandarin picture ID and picture naming at both times.

Table 2. *Mean (SD) of proficiency ratings, percent language input and output, and proportion correct on the lexical tasks as a function of group and time.*

	Younger time 1	Younger time 2	Older time 1	Older time 2
Mandarin rating[a]	4.05 (0.56)	4.23 (0.63)	3.69 (0.60)	3.79 (0.69)
English rating[a]	3.63 (0.88)	4.40 (0.64)	4.40 (0.50)	4.65 (0.40)
English input	53% (18%)	64% (11%)	61% (13%)	67% (11%)
English output	60% (22%)	70% (14%)	80% (17%)	80% (15%)
English ID	0.71 (0.21)	0.91 (0.04)	0.93 (0.10)	0.98 (0.02)
English naming	0.45 (0.24)	0.71 (0.12)	0.83 (0.10)	0.90 (0.04)
Mandarin ID	0.70 (0.10)	0.70 (0.16)	0.75 (0.13)	0.74 (0.15)
Mandarin naming	0.33 (0.13)	0.36 (0.17)	0.36 (0.20)	0.35 (0.18)

[a] Ratings were based on a five-point scale: 1 = low proficiency, 5 = high proficiency.

According to parent rating, both the younger ($t = 3.21$, $df = 11$, $p < .01$) and the older ($t = 2.15$, $df = 14$, $p < .05$) children's English proficiency increased significantly from time 1 to time 2; but their Mandarin proficiency stayed the same, $ps > .10$. The two age groups received comparable Mandarin ratings at both times and comparable English ratings at time 2, $ps > .05$. English ratings at time 1 was significantly higher for the older children, $p < .01$ (see Table 2 for means). As part of the questionnaire, parents were also asked to rate their own English proficiency using a five-point scale (0 = none, 1 = poor, 2 = fair, 3 = good, 4 = excellent). On average, the father's English proficiency was 3.37 (SD = 0.69, range = 2–4), and the mother's English proficiency was 2.67 (SD = 0.83, range = 1–4).

A face-to-face interview was conducted with the parent at both times to document the child's hour-by-hour language use. Specifically, we asked the parent to describe the child's schedule on a typical weekday and a typical weekend. For each waking hour, we asked the parent what activities the child participates in, whom the child interacts with, and what language(s) are used for communication between the child and his/her conversational partner(s). This information was later entered into an excel spreadsheet with built-in formula to calculate the child's percentage of input (amount of time hearing a language) and output (amount of time speaking a language). The algorithm takes into account all waking hours of the child's week and documents the language used for all recurring activities (e.g., two hours of Sunday Chinese school are entered as time hearing and speaking Mandarin; one hour of Saturday piano class conducted in English is entered as time hearing and speaking English).

For the younger group, there was a significant increase in both English input ($t = 2.77$, $df = 11$, $p = .02$) and English output ($t = 2.85$, $df = 11$, $p = .02$) from time 1 to time 2; for the older children, the amount of input and output was comparable at both times, $ps > .10$. Amount of English output was significantly higher for the older than younger group at time 1 ($t = 2.66$, $df = 25$, $p = .01$). The two groups had comparable amount of English input at both times and comparable English output at time 2, $ps > .05$. Five of the 12 younger children and 12 of 15 the older children went to a Sunday Chinese school that provided two hours of instruction on Chinese reading and writing.[3]

Stimuli and procedures

A picture identification task and a picture naming task were used to measure, respectively, receptive and expressive vocabulary. Target items were compiled by consulting published studies of young children's vocabulary development (Hao, Shu, Xing & Li, 2008; Kan & Kohnert, 2005; Morrison, Chappell & Ellis, 1997; Tardif, Fletcher, Zhang, Liang & Zuo, 2008). The picture naming task consisted of 62 pictures. The picture identification task consisted of 65 arrays of pictures; each array included the target, a semantic foil, a phonological foil, and an unrelated foil. Different foils were used in the two languages. All pictures were black-and-white line drawing. There were three practice items in the identification task and five practice items in the naming task. At both times the two tasks were administered in Mandarin and in English on two different days by a native speaker of each language. Order of the task and language of testing were counterbalanced across participants. Details about stimulus construction and task administration as well as the full set of stimuli can be found in Sheng et al. (2011). Task reliability (Cronbach's alpha) at time 1 and time 2 was, respectively, .84 and .91

[3] The percentage of children attending Sunday Chinese schools was higher in the older than the younger age group (80% vs. 42%), a reflection of discrepancy in age and school readiness of the two groups. As the results will illustrate, access to informal Chinese instruction did not lead to gains in Mandarin vocabulary on the part of the older children, suggesting that this level of Chinese schooling may not be sufficient to cause positive gains.

for the Mandarin picture identification task; .96 and .72 for the English picture identification task; .93 and .93 for the Mandarin picture naming task; and .97 and .91 for the English picture naming task.

Scoring and reliability

For the picture identification task, we tallied the number of correct responses as well as the number of semantic, phonological, and unrelated errors at both time points. For the picture naming task, we calculated the number of correct responses and coded errors of naming into categories using a coding scheme adapted from Sheng and McGregor (2010). Responses deviating from the lexical targets were coded into subcategories such as semantic, phonological, visual, novel words, unrelated, "don't know", and language-switches. Examples of errors are presented below. In all examples, the target is presented first and the child's response is presented after the slash.

A semantic error could be (i) a superordinate (e.g., *doctor*/*man*, *diamond*/*jewel*), in which the participant used a general term to substitute for a more specific word; (ii) a coordinate (e.g., *tiger*/*cheetah*, *comb*/*brush*); or (iii) a more specific subordinate term (e.g., *eagle*/*bald eagle*, *bear*/*polar bear*). Together, these three error types reflected taxonomic knowledge as they related to the target at the superordinate, basic, or subordinate levels of the taxonomy. Semantic errors could also be (iv) a thematically related term (e.g., *cloud*/*sky*, *diamond*/*ring*), or (v) a circumlocution (e.g., *bib*/*thingy for my slobber*, *ambulance*/*rescue truck*). Phonological errors included responses that were related to the target on a purely phonologically basis (e.g., *clown*/*clone*, *violin*/*violent*); errors that bore both semantic and phonological relations to the target (e.g., *toothbrush*/*brush*) were counted as semantic. Visual errors included (i) visual misperceptions (e.g., *diamond*/*light*, *bib*/*surf board*) due to overall visual (but not semantic) similarity; and (ii) visual misinterpretations, which may be a description of an aspect of the visual scene (e.g., *tie*/*neck*, *zoo*/*giraffe*) without attempting the actual target. Novel words included made-up compound words that described the features or functions of the targets (e.g., 海马/水马 "seahorse/water horse", 指南针/方向表 "compass/direction watch"). Unrelated errors were those that bore no semantic, phonological, or visual relations to the target (e.g., *bib*/*game*, *swan*/*kin*). Language-switched responses were noted and coded as correct (e.g., 天鹅 "swan"/*swan*), semantic (e.g., 鸵鸟 "ostrich"/*flamingo*), phonological (e.g., 蜗牛 "snail"/*nail*), unrelated (e.g., 加油站 "gas station"/*casher*) or visual (e.g., 指南针 "compass"/*clock*).

Two proficient bilinguals coded all errors independently. Point-to-point agreement averaged at 95% and ranged from 93% to 97% by sample.

Results

Accuracy

Mean accuracy scores on the four tasks at the two time points are shown in Table 2. Four two-way ANOVAs were performed to examine the effects of age (younger, older), and time (time 1, time 2) on picture identification and picture naming performance in the two languages, respectively. Significant interactions were followed with Tukey Unequal N Honest Significant Difference posthoc tests. Because the picture identification and picture naming tasks have different demands (Oller, Jarmulowicz, Pearson & Cobo-Lewis, 2011) and different psychometric properties (i.e., picture identification was subject to chance performance but picture naming was not), a direct comparison of the two was not conducted. The two languages were also not directly compared because stimulus difficulty in English and Mandarin may not be the same.

For English picture identification, the main effects of age ($F(1,25) = 17.18$, $p < .001$, $\eta_p^2 = .41$) and time ($F(1,25) = 20.50$, $p < .001$, $\eta_p^2 = .45$), and the interaction between age and time ($F(1,25) = 7.02$, $p = .013$, $\eta_p^2 = .22$) were significant. The interaction reflected a significantly higher performance at time 2 than time 1 for the younger ($p < .001$) but not the older group as well as a significant older group advantage over the younger group at time 1 ($p < .001$) but not time 2.

For English picture naming, the main effects of age ($F(1,25) = 36.60$, $p < .001$, $\eta_p^2 = .59$) and time ($F(1,25) = 47.39$, $p < .001$, $\eta_p^2 = .65$), and the interaction between age and time ($F(1,25) = 15.60$, $p < .001$, $\eta_p^2 = .38$) were significant. The interaction reflected a significantly higher performance at time 2 than time 1for the younger ($p < .001$) but not the older group. Nevertheless, the older group advantage over the younger group was significant at both times (ps $< .01$).

For Mandarin picture identification, there were no significant main effects or interaction, Fs < 1, ps $> .10$. Similarly, the effects of age, time and their interaction were not significant for Mandarin picture naming, Fs < 1, ps $> .10$. To summarize, for both picture identification and naming tasks, gains in lexical ability were significant only in English and among the younger children.

Although the two tasks and two languages were not directly compared, there appeared to be a greater modality (receptive/picture identification as opposed to expressive/picture naming) difference in Mandarin (.37) than in English (.15; means were averaged across age groups).

Tables 3 and 4 present individual children's performance profiles in the two languages. Of the 27 children, six showed increase on all four measures (#2, 6, 7, 10, 12, 24), 12 showed increase on three measures (#1, 3, 4,

Table 3. *English accuracy (in proportion) and change in individual performance from time 1 to time 2.*

	Picture ID			Picture naming		
Participant	Time 1	Time 2	Difference	Time 1	Time 2	Difference
1	0.57	0.85	0.28	0.19	0.61	0.42
2	0.48	0.94	0.46	0.32	0.73	0.40
3	0.34	0.88	0.54	0.05	0.55	0.50
4	0.38	0.85	0.46	0.21	0.68	0.47
5	0.71	0.94	0.23	0.39	0.63	0.24
6	0.78	0.94	0.15	0.47	0.55	0.08
7*	0.82	0.88	0.06	0.45	0.65	0.19
8*	0.92	0.95	0.03	0.79	0.89	0.10
9	0.92	0.94	0.02	0.82	0.85	0.03
10*	0.85	0.97	0.12	0.65	0.85	0.21
11*	0.85	0.89	0.05	0.60	0.74	0.15
12*	0.89	0.94	0.05	0.50	0.76	0.26
13*	0.95	0.98	0.03	0.89	0.90	0.02
14*	0.98	0.97	−0.02	0.82	0.89	0.06
15*	0.92	1	0.08	0.76	0.90	0.15
16*	0.91	0.95	0.05	0.85	0.84	−0.02
17*	0.91	0.95	0.05	0.77	0.94	0.16
18	0.92	1	0.08	0.81	0.85	0.05
19	0.95	0.97	0.02	0.89	0.94	0.05
20*	0.97	0.97	0	0.87	0.97	0.10
21*	0.57	0.97	0.40	0.53	0.82	0.29
22	0.97	0.98	0.02	0.94	0.94	0
23*	0.92	1	0.08	0.81	0.89	0.08
24*	0.97	0.98	0.02	0.85	0.90	0.05
25*	1	0.98	–0.02	0.92	0.95	0.03
26*	0.97	1	0.03	0.85	0.87	0.02
27*	1	1	0	0.92	0.92	0

* These children went to Sunday Chinese schools that offered approximately two hours of instruction per week on Chinese literacy skills.
Note: Positive difference score indicates more pictures were identified or named at time 2 than time 1. Negative difference score indicates fewer pictures were identified or named at time 2. Participants #1–#12 belong to the younger age group and participants #13–#27 belong to the older age group.

5, 8, 11, 13, 14, 18, 21, 25, 26), six showed increase on two measures (#9, 15, 17, 19, 20, 23), and three (#16, 22, 27) showed increase on only one measure and decrease/no change on the other three measures. In particular, child #27, the oldest participant, scored 100% on English identification at both times and 92% on English naming at both times. This child also showed a 9% decrease for Mandarin identification and an 11% increase for Mandarin naming. In contrast, child #2, the second youngest participant, showed an increase of 46%, 40%, 4.6%, and 6.5% on English identification, English naming, Mandarin identification, and Mandarin naming, respectively. Child #14 showed a 1.5% decrease in English identification and an increase of 6.5%, 14%, and 29% on English naming, Mandarin identification, and Mandarin naming. This child's time 2 testing coincided with the family's return from a summer vacation in Taiwan. Overall, there was a trend for the younger children to show more increases on the dependent measures than the older children.

We conducted two chi-square analyses to compare across languages the number of children who showed one of the three possible change patterns in accuracy level (i.e., increase, decrease, or no change over time) for the picture identification and naming tasks, respectively. Age group was not included as an independent variable due to the unequal size of the two age groups. There was a significant interaction between test language and change pattern for the identification task, $\chi^2 = 9.92$, $df = 2$, $p = .007$, effect size (Cramer's V) = .43; and the naming task, $\chi^2 = 12.27$, $df = 2$, $p = .002$, Cramer's V = .48. For both tasks, the number of children who showed a

https://doi.org/10.1017/S1366728913000746 Published online by Cambridge University Press

Table 4. *Mandarin accuracy (in proportion) and change in individual performance from time 1 to time 2.*

Participant	Picture ID			Picture naming		
	Time 1	Time 2	Difference	Time 1	Time 2	Difference
1	0.68	0.78	0.11	0.40	0.37	−0.03
2	0.86	0.91	0.05	0.58	0.65	0.06
3	0.65	0.65	0	0.32	0.15	−0.18
4	0.69	0.65	−0.05	0.37	0.39	0.02
5	0.57	0.52	−0.05	0.11	0.15	0.03
6	0.89	0.91	0.02	0.45	0.61	0.16
7*	0.65	0.75	0.11	0.16	0.44	0.27
8*	0.78	0.69	−0.09	0.27	0.39	0.11
9	0.62	0.31	−0.31	0.24	0.15	−0.10
10*	0.63	0.71	0.08	0.39	0.44	0.05
11*	0.66	0.69	0.03	0.24	0.23	−0.02
12*	0.72	0.78	0.06	0.37	0.42	0.05
13*	0.88	0.92	0.05	0.53	0.44	−0.10
14*	0.60	0.74	0.14	0.08	0.37	0.29
15*	0.66	0.63	−0.03	0.27	0.16	−0.11
16*	0.57	0.49	−0.08	0.06	0.05	−0.02
17*	0.80	0.80	0	0.34	0.29	−0.05
18	0.72	0.83	0.11	0.47	0.47	0
19	0.91	0.78	−0.12	0.39	0.26	−0.13
20*	0.77	0.72	−0.05	0.24	0.26	0.02
21*	0.80	0.86	0.06	0.65	0.50	−0.15
22	0.55	0.43	−0.12	0.42	0.23	−0.19
23*	0.74	0.71	−0.03	0.34	0.23	−0.11
24*	0.60	0.63	0.03	0.11	0.35	0.24
25*	0.92	0.94	0.02	0.76	0.81	0.05
26*	0.94	0.89	−0.05	0.44	0.47	0.03
27*	0.83	0.74	−0.09	0.31	0.42	0.11

* These children went to Sunday Chinese schools that offered approximately two hours of instruction per week on Chinese literacy skills.
Note: Positive difference score indicates more pictures were identified or named at time 2 than time 1. Negative difference score indicates fewer pictures were identified or named at time 2. Participants #1–#12 belong to the younger age group and participants #13–#27 belong to the older age group.

decrease was significantly lower than expected in English but significantly higher than expected in Mandarin.

Correlational analyses were conducted to examine which of the 10 language background factors (age at time 1, amount of cumulative English experience, English proficiency rating at time 1 and time 2, Mandarin proficiency rating at time 1 and time 2, input at time 1 and time 2, output at time 1 and time 2) were related to changes in lexical access over time. Because the number of months between the two testing varied across participants, time interval was also included in the correlational analyses to see if this factor was related to changes in lexical access. An adjusted p level of .0023 was used to account for the number of correlations conducted per language (.05 divided by 22). As shown in Table 5, significant correlations were identified in English only. Specifically, greater increase in picture identification and picture naming accuracy was associated with younger age, less cumulative English exposure, lower English rating at time 1, lower English output at time 1, and higher Mandarin rating at time 2 (for naming only).

Errors

Picture identification

We computed the average proportions of picture identification errors by dividing the number of each type of errors by the number of total errors in each child and averaging over the 27 children. Proportions are used to take into account the differences in raw number of errors

https://doi.org/10.1017/S1366728913000746 Published online by Cambridge University Press

Table 5. *Correlation coefficients between parental rating, language environment variables, and difference scores in lexical performance from time 1 to time 2. ID = picture identification; naming = picture naming.*

	English		*Mandarin*	
	ID	*naming*	*ID*	*naming*
Age time 1	−.58***	−.70***	−.16	−.10
Years of English exposure	−.57***	−.61***	.14	.00
English rating time 1	−.78***	−.71***	−.29	.12
English rating time 2	−.09	−.01	−.31	−.17
Mandarin rating time 1	.43	.41	.26	.10
Mandarin rating time 2	.56	.61***	.47	.22
English input time 1	−.42	−.29	−.30	−.14
English input time 2	−.22	−.14	−.00	.12
English output time 1	−.65***	−.62***	−.38	.02
English output time 2	−.42	−.38	−.32	−.19
Test interval[a]	.05	.15	.35	−.09

*** $p < .0023$
[a] Number of months between time 1 and time 2.

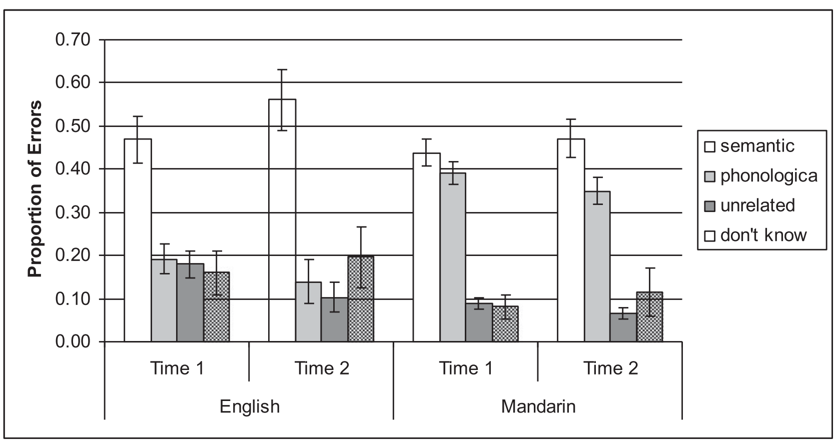

Figure 1. Mean proportion of errors on the picture identification task as a function of error type, language, and time. Bars denote standard errors.

across languages. Across the 27 children, there were a total of 106 semantic errors, 83 phonological errors, 59 unrelated errors, and 45 “don’t know” errors for English time 1; and these numbers decreased to 45, 13, 14, and 14 for English time 2. The raw numbers of semantic, phonological, unrelated, and “don’t know” errors were, respectively, 190, 193, 48, and 43 for Mandarin time 1; and 189, 169, 36, and 95 for Mandarin time 2. The proportion of various types of picture identification errors within a language cannot be directly compared due to dependence of variance. Nevertheless, inspection of Figure 1 revealed that in English, semantic errors predominated (averaging 47% of all errors at time 1 and 56% at time 2) over the other three error types, which were comparable in proportion. In Mandarin, at time 1, semantic and phonological errors were similarly common and outnumbered unrelated and “don’t know” errors; at time 2, semantic errors had a 12% lead over phonological errors and both were more common than unrelated and “don’t know” errors.

To examine if picture identification errors differed by age, language, and time, the proportion of semantic, phonological, unrelated, and “don’t know” errors were plotted and patterns were identified through visual inspection of means and standard errors. Emerging patterns were then confirmed through t-tests. Two trends were noted in the data (see Figure 1). First, there were proportionally more phonological errors for Mandarin ($M = .37$, $SE = .026$) than for English ($M = .15$, $SE = .025$). This was confirmed by t-test, $t = 6.37$, $df = 26$, $p < .001$. Second, there were proportionally more unrelated errors at time 1 ($M = .13$, $SE = .018$) than time 2 ($M = .07$, $SE = .018$), as confirmed by t-test, $t = 2.55$, $df = 26$,

https://doi.org/10.1017/S1366728913000746 Published online by Cambridge University Press

Table 6. *Mean proportions (standard deviations) of each error type in the picture naming task in time 1 and time 2.*

	English		Mandarin	
Error type	Time 1	Time 2	Time 1	Time 2
Semantic	0.491 (0.231)	0.644 (0.227)	0.216 (0.178)	0.235 (0.163)
Phonological	0.032 (0.070)	0.004 (0.016)	0.013 (0.027)	0.022 (0.035)
Unrelated	0.066 (0.091)	0.015 (0.041)	0.018 (0.024)	0.011 (0.016)
"Don't know"	0.333 (0.299)	0.239 (0.248)	0.495 (0.317)	0.608 (0.281)
Visual	0.063 (0.093)	0.097 (0.118)	0.026 (0.037)	0.026 (0.040)
Novel words	0	0	0.005 (0.016)	0.009 (0.017)
Language-switch				
Correct	0.011 (0.043)	0	0.175 (0.214)	0.067 (0.151)
Semantic	0.001 (0.006)	0	0.036 (0.051)	0.021 (0.055)
Phonological	0	0	0.002 (0.008)	0
Unrelated	0	0	0.007 (0.014)	0
Visual	0.003 (0.018)	0	0.007 (0.016)	0.001 (0.004)
Total	0.015 (0.066)	0	0.227 (0.276)	0.089 (0.197)

$p = .02$. Visual inspection and focused *t*-tests did not reveal any noticeable patterns in the occurrence of semantic and "don't know" errors.

Picture naming

The average proportions of picture naming errors are presented in Table 6. Across the 27 children, there were a total of 561 errors for English time 1, 310 errors for English time 2, 1094 errors for Mandarin time 1, and 1075 errors for Mandarin time 2. Although statistical comparison of error distribution within a language cannot be conducted, inspection of Table 6 indicated that for English, semantic errors were the dominant error type (time 1: 49%; time 2: 64%) at both times. "Don't know" errors were the second most common (time 1: 33%; time 2: 24%). Phonological, unrelated, and other (i.e., visual, language-switches) errors were less common, and novel word errors were non-existent. For Mandarin, semantic errors (22%) and language-switches (23%) were equally common and both were less common than "don't know" errors (50%) at time 1; at time 2, "don't know" errors increased to 61%, semantic errors stayed at the same level (24%), and language-switches became less common (9%). Phonological, visual, unrelated, and novel words errors were rare in Mandarin at both times.

To examine if the production of the more frequent subtypes of picture naming errors were different across age groups, time, and languages, the proportion of semantic, "don't know", and language-switch errors were plotted (see Figure 2) and patterns were identified through visual inspection of means and standard errors. Emerging patterns were then confirmed through *t*-tests.

For semantic errors, there were two trends. First, comparison between the upper and lower panels of Figure 2 suggested a higher proportion of English semantic errors in the older ($M = .62$, $SE = .056$) than the younger children ($M = .50$, $SE = .056$); however, this difference did not reach significance, $p = .14$. Second, when the two age groups were collapsed, there was a trend for the proportion of semantic errors to increase from time 1 ($M = .49$, $SE = .045$) to time 2 ($M = .64$, $SE = .044$) in English $t = 4.53$, $df = 26$, $p < .001$, whereas these numbers did not change over time in Mandarin (time 1: $M = .22$, $SE = .034$; time 2: $M = .23$, $SE = .031$), $p > .50$.

For "don't know" errors, there were two trends. First, comparison between the upper and lower panels of Figure 2 suggested that in Mandarin, the older children ($M = .65$, $SE = .055$) made proportionally more "don't know" errors than the younger children ($M = .43$, $SE = .071$). This was confirmed by a *t*-test, $t = 2.48$, $df = 25$, $p = .02$. The reverse was true in English, although the difference was not statistically significant (younger: $M = .37$, $SE = .065$; older: $M = .22$, $SE = .071$), $p = .16$. Second, when the two age groups were collapsed, there was a trend for a significant decrease in "don't know" errors for English from time 1 ($M = .33$, $SE = .057$) to time 2 ($M = .24$, $SE = .048$), $t = 2.80$, $df = 26$, $p = .009$, whereas in Mandarin there was a sizable but nonsignificant increase in "don't know" errors from time 1 ($M = .49$, $SE = .061$) to time 2 ($M = .61$, $SE = .054$), $p = .09$.

When tested in English, language-switches were almost non-existent. In Mandarin, when the two age groups were combined, children were more likely to

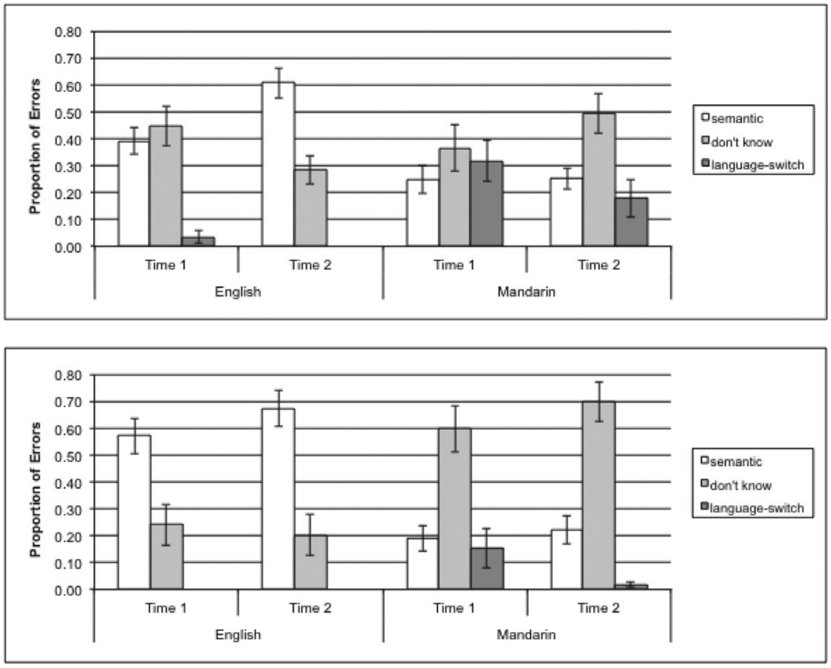

Figure 2. Mean proportion of errors on the picture naming task in the younger children (upper panel) and older children (lower panel) as a function of error type, language, and time. Bars denote standard errors.

switch to English at time 1 ($M = .23$, $SE = .053$) than time 2 ($M = .09$, $SE = .038$), $t = 2.74$, $df = 26$, $p = .01$. These patterns are shown in Figure 2.

To summarize, analyses of naming error distribution indicated age differences in English such that the older children made proportionally more semantic errors and fewer "don't know" errors than the younger children in this language. Effect of time was manifested differently by language. In English, semantic errors increased but "don't know" errors decreased. In Mandarin, semantic errors were stable, "don't know" errors increased, and language-switches decreased.

Discussion

This study aims to enhance our understanding of lexical-semantic skill development during a dynamic transitional period in the lives of young bilingual children who are becoming increasingly English-dominant. We analyzed accuracy at the group and individual levels and examined error types to quantify and qualify the performance changes associated with age, language, and task.

With regard to the first research question, "What is the profile of L1 and L2 vocabulary development over time and how does age affect development?", the current longitudinal study yielded findings that are consistent with our previous cross-sectional investigation (Sheng et al., 2011). That is, bilingual children showed gains in English vocabulary and no gains in Mandarin vocabulary. Lack of gains in Mandarin occurred even though children were experiencing an average of about 34% Mandarin input at time 2 (33% for younger and 35% for older age group). Gains in English were greater in the younger children and statistically non-significant in the older children. This latter finding may be an artifact of the current English measures, which may have been too easy for some of the oldest children. Also similarly to Sheng et al. (2011), the performance gap between the identification and naming task was much wider for Mandarin than English. These findings indicated a shift toward greater proficiency in English during the nearly 16-month interval.

Although there was no systematic decline in L1 lexical-semantic skills, the lack of L1 growth among the preschool to early school-age children in the present study was in contrast to findings observed for Spanish-speaking children, who continued to show significant growth in L1 vocabulary even during school-age years (Cobo-Lewis et al., 2002; Kohnert, 2002; Uchikoshi, published online September 11, 2012). Recall that Uchikoshi (published online September 11, 2012) also included samples of Cantonese-speaking children from low-income families. Between kindergarten and second grade, the Cantonese-speaking children who went to bilingual schools showed an 11-point increase in L1 receptive vocabulary; those who went to English mainstream schools showed an increase of seven points in L1 receptive vocabulary. In contrast, the current sample of Mandarin-speaking children, who came from middle SES families and went to English mainstream schools, showed a 1% decrease (when averaged across ages) in L1 receptive vocabulary

https://doi.org/10.1017/S1366728913000746 Published online by Cambridge University Press

over the 16 months period. It is important to note that unlike the children in Uchikoshi's study, who lived in Northern Californian communities with high percentages of L1 speakers, the children in the current study resided in a Southern city with relatively low representation of L1 speakers. Also unlike the children in Uchikoshi (published online September 11, 2012), none of the current children received L1 instruction during regular school hours. The only form of semi-formal L1 instruction took place at Sunday Chinese schools that lasted for two hours a week. Taken together, the discrepancies between the current study and previous studies suggest that linguistic typology, degree of community L1 support, family characteristics, and access to L1 schooling may have conspired to yield a pattern of fossilized L1 vocabulary (Goldenberg et al., 2011; Han, 2008; Marinova-Todd & Uchikoshi, 2011).

However, these group patterns belie the complexity in bilingual vocabulary acquisition at the individual level. Examination of individual's change patterns indicated great inter- and intra-individual variability in children's profile. As predicted, only a minority (22%) of the children showed increases on all four measures. However, all children showed increases on at least one measure and many children (44%) showed increases on three measures. For some children, particularly the older ones, increases on the English identification measure were rather small. This was in part due to limitations of the measurement scale. A large number of children (two younger and 14 older) were above 90% accurate at time 1 on the English identification task. This level of accuracy may have left little room for improvement. Nevertheless, this limitation did not prevent us from seeing an English advantage in terms of the direction of developmental changes as a majority of the children (i.e., 85% for picture ID and 89% for picture naming showed increases on the English measures but only half of them showed increases on the Mandarin measures (i.e., 48% for picture ID and 52% for picture naming).

Correlation analyses further informed us about the mechanisms of change in lexical-semantic skills. Results indicated that younger children and children who had lower English proficiency and/or lower English output at time 1 made greater gains on the two English lexical tasks. Given that our sample consisted mainly of early sequential bilinguals, these three variables – younger age, lower initial English proficiency, and lower initial English output, were most likely correlated with each other. These findings are consistent with Goldenberg et al. (2011), who suggested a greater potential for growth when prior lexical knowledge was low. At the same time, Mandarin difference scores were not significantly correlated with any of the background variables. This lack of correlations for L1 measures was not without precedence: Yan and Nicoladis (2009) also found that age was positively correlated with French–English bilingual children's receptive vocabulary score and naming performance in their stronger language (English), but not in their weaker language (French). Factors other than those measured in the current study may be at play in explaining vocabulary changes in the weaker language. For instance, child #14, who demonstrated robust gains in Mandarin scores, was tested shortly after the return from a lengthy trip in Taiwan, where Mandarin is the official language. Immersion in an L1 environment may have enabled substantial L1 vocabulary learning or re-activated L1 words that are latent in the child's lexicon (Oller et al., 2011). Finally, the positive correlation between time 2 Mandarin proficiency rating and English vocabulary growth suggests the use of a crosslinguistic semantic bootstrapping mechanism among children who were able to maintain their L1 proficiency. Bilingual lexical development does not have to be subtractive even among children who are rapidly shifting towards L2 dominance. Future studies with larger samples may identify subgroups of children who are or are not able to build their dual language lexicon and illuminate the factors associated with these distinct growth profiles.

Our last research question pertains to language and time effects on error patterns. Errors on lexical-semantic tasks are under-studied but they can provide a window into linguistic and cognitive changes associated with development (Jaeger, 2005). With regard to picture identification errors, the current study revealed two cross-language differences. First, in Mandarin picture identification, semantic and phonological errors were roughly comparable; but in English, semantic errors were clearly dominant. Second, children made more phonological errors in Mandarin than English. These patterns were consistent with our predictions. The predominance of semantic errors in English identification tasks suggested that participants were able to reach the correct semantic neighborhood for the English word targets. At the same time, phonological errors indicated failures in accessing semantic/conceptual representations of the targets. The high frequency of phonological errors in the Mandarin identification tasks suggests that children may not have enough knowledge about the target to differentiate it from a similar-sounding word. For instance, the target 蝎子-*xie1zi3* "scorpion" and its phonological distractor 鞋子-*xie2zi3* "shoe" differ by only the tone of the first syllable. A child who has not yet mapped the word form *xie1zi3* with the correct semantic representation may be easily tripped by the phonological similarity between the target and its phonological neighbor. On the other hand, the low occurrence of phonological errors in English suggests that children may possess enough knowledge of the target (e.g., *barn*) and the phonological neighbor (e.g., *Barney*) to not choose the phonological distractor. Instead, semantic (e.g., *stable*) distractors were more likely to be chosen.

With regard to time effect on picture identification, the only significant time-related difference was for unrelated errors to show a decrease over time. Inspection of Figure 1 suggests the decrease in unrelated errors was mainly attributed to English ($p = .08$) and a simultaneous (but non-significant) increase in English semantic errors. This trend provides further evidence of growth in English lexical-semantic skills.

Error analyses in picture naming indicated age group and cross-language differences. In terms of age, the older children tended to make a higher proportion of semantic errors and a lower proportion of "don't know" errors than the younger children. Not only were the older children more accurate in English naming; when they failed to provide the correct name, the errors these children made were also closer to the targets. In terms of cross-language differences, three trends in naming error distribution were suggestive. First, in English, semantic errors predominated but in Mandarin, "don't know" errors were the most common. Second, children made proportionally more semantic errors in English than Mandarin and more "don't know" errors in Mandarin than English. Although not directly compared, language-switches were also much more prevalent in Mandarin than English. Third, in English, there was an increase in semantic errors over time and a simultaneous decrease in "don't know" errors. In Mandarin, there was no change in semantic errors, an increase in "don't know" errors, and a decrease in language-switched responses. A majority of the language-switched responses at both times were successful, meaning that the responses resulted in the correct labels in English.

These patterns were again consistent with our predictions. Semantic errors occur when children have knowledge of the target's semantic category, thematic context of use, or functional and physical features (German & Newman, 2004; McGregor, 1997; Sheng & McGregor, 2010; Yan & Nicoladis, 2009). Semantic substitutions and circumlocutions are effective communicative strategies individuals use to compensate for lexical gaps or temporary retrieval difficulties (Dörnyei & Kormos, 1998; Dörnyei & Scott, 1997; Greene et al., 2012). The frequent production of semantic substitutions in the older children reflects more advanced cognitive and linguistic skills in these children. Similarly, when tested in English children not only made fewer errors overall, the errors they made were also more informative and closer to the targets (or at least indicated access of the right semantic space).

When tested in Mandarin, the majority of naming errors were indeterminate in nature. While "don't know" errors were uninformative as for the locus of word retrieval difficulties, the decrease of language-switches and the concurrent increase of "don't know" responses suggest changes at the pragmatic and/or cognitive level. Language-switches are yet another communicative strategy children use to demonstrate their linguistic knowledge (Greene et al., 2012). Frequent and successful language-switching from L1 to English has been observed in previous studies of Spanish-speaking children (Sheng et al., 2013; Sheng, Peña, Bedore & Fiestas, 2012). While L1 to L2 switching may be a result of the children's greater English proficiency, there are additional cognitive and pragmatic explanations. Even the youngest children in the current study had the pragmatic understanding that the Mandarin-speaking examiner would most likely understand the English responses. At the same time, children may be unwilling to switch from English to Mandarin either due to strong suppression of the L1 in an L2 environment (Oller et al., 2011), preference for English during school-like activities, or the understanding that the Caucasian examiner was unlikely to understand Mandarin responses. Finally, the decrease in L1-to-English language-switches over time suggests better inhibitory control (suppression of the non-target language) whereas the increase in "don't know" responses suggests increased awareness of one's own state of L1 lexical knowledge. These explanations are speculative and need to be corroborated with concurrent measures of pragmatic abilities and cognitive control.

The current study included a relatively small and homogeneous sample of bilingual children from middle class background. In addition, the two age groups each contained small numbers of participants of a relatively wide age range. These limitations have constrained the statistical approach and the generalization of the current findings to the larger population. Future larger-scale studies with more diverse samples may uncover different patterns. Future studies should also systematically investigate the joint influences of the community, family, and school in shaping bilingual children's language growth (Goldenberg et al., 2011).

Despite these limitations, the current study contributed several new findings to the literature on bilingual children's vocabulary development. As a group, Mandarin–English bilingual children who received minimal systematic L1 instructional support showed fast expansion of English vocabulary alongside little changes in Mandarin vocabulary over a 16 months period. At the individual level, children demonstrated a variety of profiles with most children showing increases in some but not all measures. Robust growth in Mandarin vocabulary was observed in some children but the mechanism for L1 vocabulary growth appeared to be related to factors not under current investigation (e.g., recent L1 immersion experience). Accelerated English growth was associated with lower initial English knowledge. Changes in bilingual children's error types suggest effective use of communicative strategies, increased L2 semantic knowledge, and potentially enhanced pragmatic understanding of the communicative context.

References

Barac, R., & Bialystok, E. (2012). Bilingual effects on cognitive and linguistic development: Role of language, cultural background, and education. *Child Development, 83,* 413–422.

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

Cobo-Lewis, A. B., Pearson, B. Z., Eilers, R. E., & Umbel, V. C., (2002). Effects of bilingualism and bilingual education on oral and written English skills: A multifactor study of standardized test outcomes. In D. K. Oller & R. E. Eilers (eds.), *Language and literacy in bilingual children*, pp. 64–97. Buffalo, NY: Multilingual Matters.

Dörnyei, Z., & Kormos, J. (1998). Problem-solving mechanisms in second language communication. *Studies in Second Language Acquisition, 20,* 349–385.

Dörnyei, Z., & Scott, M. L. (1997). Communication strategies in a second language: Definitions and taxonomies. *Language Learning, 47,* 173–210.

Durgunoğlu, A. Y., & C. Goldenberg (eds.) (2011). *Language and literacy development in bilingual settings*. New York: The Guilford Press.

German, D. J., & Newman, R. S. (2004). The impact of lexical factors on children's word-finding errors. *Journal of Speech, Language, and Hearing Research, 47,* 624–636.

Goldenberg, C., Reese, L., & Rezaei, A. (2011). Contexts for language and literacy development among dual-language learners. In Durgunoğlu & Goldenberg (eds.), pp. 3–25.

Greene, K., Bedore, L. M., & Peña, E. D. (2012). *Bilingual children's lexical strategy selection under different narrative conditions*. Poster presented at the 33rd Symposium on Research in Child Language Disorders, Madison, WI.

Gutiérrez-Clellen, V. F., & Kreiter, J. (2003). Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics, 24,* 267–288.

Hammer, C. S., Lawrence, F. R., & Miccio, A. W. (2008). Exposure to English before and after entry to Head Start: Bilingual children's receptive language growth in Spanish and English. *International Journal of Bilingual Education and Bilingualism, 11,* 30–56.

Han, W.-J. (2008). The academic trajectories of children of immigrants and their school environments. *Developmental Psychology, 44,* 1572–1590.

Hao, M., Shu, H., Xing, A., & Li, P. (2008). Early vocabulary inventory for Mandarin Chinese. *Behavior Research Methods, 40,* 728–733.

Jaeger, J. J. (2005). *Kids' slips: What young children's slips of the tongue reveal about language development*. Mahwah, NJ: Lawrence Erlbaum.

Kan, P. F., & Kohnert, K. (2005). Preschoolers learning Hmong and English: Lexical-semantic skills in L1 and L2. *Journal of Speech, Language, and Hearing Research, 48,* 372–383.

Kohnert, K. (2002). Picture naming in early sequential bilinguals: A 1-year follow-up. *Journal of Speech, Language, and Hearing Research, 45,* 759–771.

Kohnert, K., Bates, E., & Hernandez, A. E. (1999). Balancing bilinguals: Lexical-semantic production and cognitive processing in children learning Spanish and English. *Journal of Speech, Language, and Hearing Research, 42,* 1400–1413.

Marinova-Todd, S. H., & Uchikoshi, Y. (2011). The role of first language in oral language development in English: The case of both alphabetic and nonalphabetic languages. In Durgunoğlu & Goldenberg (eds.), pp. 29–60.

McGregor, K. K. (1997). The nature of word finding errors of preschoolers with and without word-finding deficits. *Journal of Speech and Hearing Research, 40,* 1232–1244.

Morrison, C. M., Chappell, T. D., & Ellis, A. W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *The Quarterly Journal of Experimental Psychology, 50,* 528–559.

Oller, D. K., Jarmulowicz, L, Pearson, B. Z., & Cobo-Lewis, A. B. (2011). Rapid spoken language shift in early second-language learning: The role of peers and effects son the first language. In Durgunoğlu & Goldenberg (eds.), pp. 94–120.

Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language and Learning, 43,* 93–120.

Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1995). Cross language synonyms in the lexicons of bilingual infants: One language or two? *Journal of Child Language, 22,* 345–368.

Restrepo, M. A. (1998). Identifiers of predominantly Spanish-speaking children with language impairment. *Journal of Speech, Language, and Hearing Research, 41,* 1398–1411.

Rodríguez, J., Díaz, R., Duran, D., & Espinosa, L. (1995). The impact of bilingual preschool education on the language development of Spanish-speaking children. *Early Childhood Research Quarterly, 10,* 475–490.

Sheng, L., Bedore, L. M., Peña, E. D., & Fiestas, C. E. (2013). Semantic development in Spanish–English bilingual children: Effects of age and language experience. *Child Development, 84,* 1034–1045.

Sheng, L., Lu, Y., & Kan, P.-F. (2011). Lexical development in Mandarin–English bilingual children. *Bilingualism: Language and Cognition, 14,* 579–587.

Sheng, L., & McGregor, K. K. (2010). Object and action naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 53,* 1704–1719.

Sheng, L., Peña, E. D., Bedore, L. M., & Fiestas, C. E. (2012). Semantic deficits in Spanish–English bilingual children with language impairment. *Journal of Speech, Language, and Hearing Research, 55,* 1–15.

Tardif, T., Fletcher, P., Zhang, Z. X., Liang, W. L., & Zuo, Q. H. (2008). *The Chinese Communicative Development Inventory (Putonghua and Cantonese versions): Manual, forms and norms*. Beijing: Peking University Medical Press.

Uchikoshi, Y. Development of vocabulary in Spanish-speaking and Cantonese-speaking English language learners. *Applied Psycholinguistics,* doi:10.1017/

S0142716412000264. Published online by Cambridge University Press, September 11, 2012.

Winsler, A., Díaz, R., Espinosa, L., & Rodríguez, J. (1999). When learning a second language does not mean losing the first: Bilingual language development in low income Spanish-speaking children attending bilingual preschool. *Child Development, 70,* 349–362.

Yan, S., & Nicoladis, E. (2009). Finding *le mot juste*: Differences between bilingual and monolingual children's lexical access in comprehension and production. *Bilingualism: Language and Cognition, 12,* 323–335.

https://doi.org/10.1017/S1366728913000746 Published online by Cambridge University Press