SCIENTIFIC STUDIES OF READING, *9*(1), 43–61

Bilingualism, Biliteracy, and Learning to Read: Interactions Among Languages and Writing Systems

Ellen Bialystok, Gigi Luk, and Ernest Kwan
York University

Four groups of children in first grade were compared on early literacy tasks. Children in three of the groups were bilingual, each group representing a different combination of language and writing system, and children in the fourth group were monolingual speakers of English. All the bilingual children used both languages daily and were learning to read in both languages. The children solved decoding and phonological awareness tasks, and the bilinguals completed all tasks in both languages. Initial differences between the groups in factors that contribute to early literacy were controlled in an analysis of covariance, and the results showed a general increment in reading ability for all the bilingual children but a larger advantage for children learning two alphabetic systems. Similarly, bilinguals transferred literacy skills across languages only when both languages were written in the same system. Therefore, the extent of the bilingual facilitation for early reading depends on the relation between the two languages and writing systems.

Learning to read is indisputably the premier academic achievement of early schooling. It prepares children for their educational futures and is the key to the possibilities that their futures hold for them. Thus, if knowing two languages at the time that literacy is introduced, or learning to read in a language that is not the child's dominant one, or acquiring literacy simultaneously in two languages affects the outcome of literacy instruction, then it would be important to know that. These possibilities affect a sizable portion of the world's children: A significant number are bilingual at the time they begin reading, many are instructed in a language they do not speak at home, and some number of those are expected to acquire this skill in two languages.

Requests for reprints should be sent to Ellen Bialystok, Department of Psychology, York University, 4700 Keele Street, Toronto, Ontario, M3J 1P3, Canada. E-mail: ellenb@yorku.ca

The writing system that a language uses affects children's acquisition of literacy because each system is based on a different set of symbolic relations and requires different cognitive skills (Coulmas, 1989). Alphabetic systems are based on correspondences between phonemes and graphemes; even though different scripts can be used, such as Roman or Semitic, they share the essential feature that the graphemes represent phonetic segments. Syllabaries, such as Korean, establish correspondences between consonant–vowel groups and graphemes. These relations place different demands on children's analysis of spoken language and their recording of the language in print. Character languages, such as Chinese, select the morpheme as the basic linguistic unit and associate those meaningful segments with characters indicating both semantic and some phonological properties. Because the task of learning to read in each of these writing systems is different, any effect of bilingualism on learning to read will depend on the type of writing system used in each language.

There are two reasons that literacy may proceed differently for bilingual and monolingual children. The first is that bilinguals develop several of the background skills for literacy differently from monolinguals; the second is that bilinguals may have the opportunity to transfer the skills acquired for reading in one language to reading in the other. In both cases, the relation between the writing systems in the two languages determines the commonality in the cognitive skills required for reading and may also determine the extent to which bilingualism affects literacy acquisition. This study examines the role of bilingualism in early literacy acquisition for children whose two writing systems have different relations to each other.

The first reason to expect bilingualism to influence literacy acquisition is the differential development of the prerequisite skills by monolingual and bilingual children. Three skills crucial for literacy acquisition are oral proficiency, metalinguistic awareness, and general cognitive development. Oral vocabulary has been shown to influence children's acquisition of literacy (Adams, 1990; Dale, Crain-Thoreson, & Robinson, 1995; McBride-Chang & Chang, 1995; Stahl & Fairbanks, 1986; Stanovich, 1986), but preschool bilingual children command a smaller vocabulary than comparable monolingual speakers of each language (Ben-Zeev, 1977; Bialystok, 1988; Merriman & Kutlesic, 1993; Rosenblum & Pinker, 1983; Umbel, Pearson, Fernández, & Oller, 1992). In a large-scale study of nearly 1,000 Spanish–English bilinguals in Miami, Cobo-Lewis, Pearson, Eilers, and Umbel (2002) reported superior vocabulary by monolinguals even after controlling for socioeconomic status (SES), a gap that did not close until fifth grade; the gap persisted even for bilingual children in English-speaking homes, although the size was reduced. This relation with home language replicates a result reported by Bialystok and Herman (1999) for English–French bilingual children. These differences in vocabulary level may disadvantage bilingual children in early literacy compared to their monolingual peers.

Regarding metalinguistic awareness, research has repeatedly confirmed the importance of phonological awareness for alphabetic reading (e.g., Bryant & Goswami, 1987; Liberman, Shankweiler, Liberman, Fowler, & Fischer, 1977; Morais, 1987; Perfetti, Beck, Bell, & Hughes, 1988; Wagner, Torgesen, & Rashotte, 1994) but it is also related to children's progress in learning to read nonalphabetic languages such as Chinese (Hanley, Tzeng, & Huang, 1999; Ho & Bryant, 1997; Shu, Anderson, & Wu, 2000). Relatively little research has examined the acquisition of phonological awareness for bilingual children, but studies typically report a bilingual advantage for 5-year-olds that disappears by age 6 when children begin reading instruction (Bruck & Genesee, 1995; Campbell & Sais, 1995; Yelland, Pollard, & Mercuri, 1993). In our study of monolingual and bilingual children between 5 and 7 years old, we found only limited evidence for bilingual advantages on some tasks (Bialystok, Majumder, & Martin, 2003). Moreover, the relation between the two languages is probably relevant, as some structures appear more salient in certain languages than others, influencing children's access to phonological awareness (Bruck & Genesee, 1995). In sum, however, there is no clear evidence that bilinguals are more advanced in developing phonological awareness than monolinguals.

Finally, levels of cognitive development are clearly related to literacy achievement, and some of these may also differ between monolinguals and bilinguals. Research by Geva and Siegel (2000) reported an interaction between cognitive factors such as working memory, and language- specific factors, such as orthography, that even varied within language when the orthographic demands changed, specifically, whether the voweled or unvoweled form of written Hebrew was used. These results set important limits on attempts to generalize the effect of cognitive development on reading outcomes by demonstrating the contribution of the language and how it is written. Therefore, if there are overall bilingual advantages in acquiring literacy skills, they should be limited by the language and writing systems of the two languages.

The second reason that bilingualism might alter the course of literacy acquisition is the possibility that skills can transfer to a similar domain in the other language. The majority of this literature has shown positive transfer of literacy skills across languages (e.g., Geva & Siegel, 2000; Geva, Wade-Woolley, & Shany, 1997; Oller & Eilers, 2002). Phonological awareness, a crucial basis for reading, also transfers readily and even relates to reading in the other language. As with reading, however, the extent of such transfer likely depends on the relation between the languages and the relation between the writing systems. The majority of this research has been conducted with children learning two alphabetic systems, and most of these studies have reported positive transfer of phonological skills across languages for bilinguals (Geva & Siegel, 2000; Gholamain & Geva, 1999; Wade-Woolley & Geva, 2000). Other studies have reported correlations between phonological awareness in English and either Spanish (Durgunoğlu, 1998;

Durgunoğlu, Nagy, & Hancin-Bhatt, 1993; Lindsey, Manis, & Bailey, 2003) or French (Comeau, Cormier, Grandmaison, & Lacroix, 1999) for bilingual children as well as significant influences between phonological awareness in one of these languages and word recognition in the other. In other words, the phonological awareness skills developed in one language transferred to reading ability in another language.

A stronger test of transfer of phonological awareness would be a demonstration for languages that differ in their writing system, especially when one language is not based on phonological representations. Huang and Hanley (1994) obtained significant correlations between phonological awareness skills in the two languages for Chinese–English bilinguals living in Taiwan and Hong Kong, even though phonics instruction for reading was available only to the children in Taiwan. In this case, however, there was no evidence that phonological awareness in one language influenced reading in the other. Luk (2003) reported the same pattern in a group of Chinese–English bilinguals: Phonological awareness in the two languages was significantly correlated even after controlling for working memory and nonverbal intelligence, but there was no effect across languages from phonological awareness to reading. In contrast, Gottardo, Yan, Siegel, and Wade-Woolley (2001) examined Cantonese–English bilingual children and found significant correlations between Chinese rhyme detection and English phonological and reading measures. Unlike the results of Luk, hierarchical regression analyses indicated that Chinese rhyme detection predicted English reading beyond the variance accounted for by English phoneme deletion. Gottardo et al. (2001) concluded that "phonological processing skills in a child's L1 can influence reading performance in an alphabetic orthography, regardless of the orthography used to represent in the child's L1" (p. 540). However, rhyme awareness was the only measure of phonological awareness, and these results may not extend to other dimensions of this complex skill. Moreover, the "odd-one-out" paradigm that was used carries considerable working memory demands, but working memory was not assessed, so the results are difficult to interpret. Finally, similarities in English and Chinese phonological awareness do not necessarily imply common variance in English and Chinese reading; in fact, Chinese and English reading measures did not correlate (Gottardo et al., 2001). Thus, in spite of evidence for a relation between phonological awareness abilities across the two languages, the study is less conclusive on the effect of those abilities in reading across languages.

Research on transfer of reading skills has also produced mixed results. As noted previously, there was no correlation between reading in English and Chinese for the children tested in the study by Gottardo et al. (2001). In other research, Geva and Siegel (2000) tested children from Grade 1 to Grade 5 who were English speakers attending Hebrew schools. Lists of English and Hebrew words and pseudowords were given to the children to measure their reading abilities in both languages. The researchers found that age predicted performance for real and pseudoword reading in English but not in Hebrew. Thus, children showed little improvement after they acquired the basic skills in reading Hebrew but English reading improved steadily from grade to grade. The same pattern was observed in a study of Persian–English bilinguals (Gholamain & Geva, 1999). In both cases, the difference between reading progress in the two languages was attributed to the different levels of orthographic transparency. In contrast, the study by Luk (2003) showed no correlation between reading in English and Chinese after the variance of nonverbal reasoning skills and working memory was controlled. Nonetheless, the transfer of phonological awareness across the two languages remained robust after controlling for the two nonlinguistic variables. Luk's conclusion was that phonological awareness reflected a general processing ability even when it is acquired in the context of different writing systems but reading develops individually for the specific demands of the writing system being used.

Several trends are evident in these studies. First, the pattern for the correlation of phonological skills across languages is different from that for reading skills, suggesting that the transferability of these two abilities is not the same. Second, the extent to which children transfer their skill in one language to the other language depends on the similarity of the systems, phonological structure in one case and writing system in the other. In that case, a more detailed description that takes account of these factors is needed to explain the transfer of relevant skills in the early literacy acquisition of bilingual children.

This study investigates these issues by examining decoding ability in four groups of children. Three of the groups are bilingual, but the languages and writing systems of the bilingual children have different similarity relationships. For Spanish–English bilinguals, the languages are similar (Indo-European) and both are written alphabetically in a Roman script; for Hebrew–English bilinguals, the languages are different (Indo-European vs. Semitic) and both are written alphabetically (because they are learning the voweled form of Hebrew) but use different scripts; for Chinese–English bilinguals, the languages and the writing systems share no resemblance. The purpose in selecting these groups is to isolate the role of bilingualism on children's early progress in reading and to identify the role of language and writing system similarity on that effect. There are two hypotheses. The first is that the prerequisites to literacy that develop differently in monolinguals and bilinguals will lead to an advantage for bilingual children on reading measures once these measures have been controlled but that the advantage will be greater if the two languages use the same writing system. The second is that the writing systems will determine the extent to which the skills developed in one language will transfer to the bilinguals' other language.

METHOD

Participants

The sample consisted of 132 Grade 1 children comprising four groups: English monolingual, Cantonese–English bilingual, Hebrew–English bilingual, and Spanish–English bilingual. All the children lived in the same metropolitan area in which English was the primary language, although other languages were also prevalent in the communities. The children from the monolingual, Cantonese–English bilingual, and Hebrew–English bilingual groups lived in the same neighborhood and attended the same schools. In the bilingual groups, all the children spoke English at school and another language at home. Based on parents' reports in the questionnaire, Peabody Picture Vocabulary Test (PPVT) scores (reported next), and educational experiences, the three groups of bilingual children were equivalent in their degree of bilingualism. Parents of English monolingual children were asked whether their children were exposed to a second language, and parents of bilingual children were asked questions regarding their home literacy environment. Teachers' reports indicated that all the children were free of hearing, behavioral, and reading problems.

There were 40 children in the English monolingual group (21 female and 19 male; *M* age = 81.1 months) recruited from two schools in similar neighborhoods. The experimenter confirmed the parental report by asking the children if they spoke other languages at home. These children attended public schools and received instruction only in English reading. The method of English reading instruction was the same for all the children, including those in the three bilingual groups, and employed a variety of approaches and activities but was largely phonics based.

The Cantonese bilingual group consisted of 29 children (16 female and 13 male; *M* age = 78.7 months) who attended public elementary schools where instruction was in English and were also enrolled in weekly Chinese classes that lasted 2 hr each. The program in the Cantonese school included instruction in speaking and reading, using a "look and say" method, with regular written tests of the material. Teachers would present a new character, indicate the pronunciation, and explain the vocabulary item in an example sentence. Each new character was taught as a whole without partitioning into smaller compartments. Parental reports indicated that the children spoke mainly Cantonese, the community was largely Chinese, and both languages were present in the environment, offering community services, newspapers, and magazines.

The Hebrew bilingual group consisted of 30 children (20 female and 10 male; *M* age = 80.8 months). These children attended a private day school in which Hebrew was the language of instruction for some subjects and English was used for others. Children received daily literacy instruction in both Hebrew and English primarily using a phonics-based approach. These children spoke primarily Hebrew at home because at least one parent was a native speaker of Hebrew. Reading instruction in both languages was phonologically based and emphasized associations between letters and sounds. Hebrew was introduced in the dotted script that includes vowel sounds, making the written language an alphabetic system with a shallow orthography.

The Spanish bilingual group consisted of 33 children (18 female and 15 male; *M* age = 82.7 months) who attended elementary schools in English but were additionally enrolled in an after- school Spanish program. Like the Cantonese classes, the program included instruction in spoken and written Spanish. Like the children in the other bilingual groups, these children spoke primarily Spanish at home, but English was spoken with their friends and at school.

Procedures

All the children received the same English tasks, and the bilingual children were also given adaptations of the tasks in their other language. The children were tested during midwinter of first grade. Testing for the monolingual children was completed in one session lasting about 30 min, and the bilingual children were visited twice, once for each language, separated by at least 1 week. The order of the languages used in testing bilingual children was counterbalanced. All the testing was conducted by three trained researchers who were fluent in English and one of the other languages. The English monolingual group was tested by the same researcher as the Cantonese bilingual group. The instructions for each testing session were always given in the language being tested. Prior to testing, the researcher conversed casually with the bilingual children (e.g., asked for their names, their favorite food, movies, and books, etc.) in both languages to ensure that children could communicate in these languages. Children who failed to respond and appeared not to understand were excluded from subsequent testing.

Instruments

PPVT–Revised. The PPVT–Revised (PPVT–R; Dunn & Dunn, 1981) is a standardized test for receptive vocabulary. The child was shown a page containing four pictures, one of which was named, and the child indicated the picture that corresponded to the word spoken by the experimenter. Testing continued until there were errors on six out of eight consecutive items. Raw scores were converted to standard scores according to the child's age in months by means of normalized tables.

The test exists in two parallel forms. Form M was used to test all the children in English and Form L was translated into Cantonese and Hebrew to assess children's competence in those languages. The translations were created by native speakers of those languages, but the translated test is not standardized so the scores can provide only a rough approximation of proficiency in those two languages. The Spanish bilingual group was tested by the Test de Vocabulario en Imagenes Peabody (TVIP; Dunn, Lugo, Padilla, & Dunn, 1986), a standardized Spanish version of PPVT–R. The reliability scores reported in the manual for PPVT–R are .79, and .61 for children who are 6 and 7 years old, respectively. The TVIP reports a reliability of .93 and .94 for children of those ages.

Forward and Backward Digit Span. Digit Span, derived from the Wechsler Intelligence Scale for Children–Revised (WISC–R; Wechsler, 1974), provides a measure of short-term verbal memory. In Forward Span, the experimenter read strings of digits that increased in length and the child repeated each string in the same order. There were two trials at each string length, beginning with two, and testing proceeded until the child failed to correctly repeat both trials at a particular string length. The longest string consisted of eight digits, creating a maximum score of 16. The procedure for the Backward Span was the same, except children repeated the numbers in the reverse order. Again, testing terminated when two errors were made at the same string length. The maximum score for the Backward Span test was 14, creating a total maximum score of 30. Similar tests were constructed in the other language for the bilingual groups. The numbers used in the translated versions were different from those in the English version. The bilingual groups were tested in both languages. The reliability scores reported in the manual of WISC–R for Digit Span are .76 and .84 for 6½- and 7½-year-old samples.

Phoneme Counting Task (Bialystok et al., 2003). This task was used to assess the child's phoneme segmentation ability. The Count, a character from *Sesame Street*®, was introduced, and the child was reminded that The Count likes to count things; today he would count the number of sounds in a word. A demonstration was given indicating how the sounds of a word could be "spread out" and then counted individually. Children were trained to say the word slowly, pay attention to the sounds, and move one marker chip from a pile for each sound they heard. The number of markers was counted to indicate the number of sounds in the word. Children received training with three words, varying in their number of phonemes, to learn this technique. When they understood the instructions, 10 test items were presented that ranged in the number of phonemes from two to five. Ten items is sufficient to distinguish among children's abilities because the task is difficult and children's level of understanding is apparent from the first few items. Parallel versions were created in Hebrew and Spanish by native speakers of these languages using corresponding numbers of phonemes in one-syllable words as in the English test. There was no Cantonese version because the phonemic structure of Cantonese is less accessible than it is in alphabetic languages and develops as consequence of acquiring alphabetic literacy (Cheung, Chen, Lai, Wong, & Hills, 2001). A few trial items were pilot tested on adult native speakers of Cantonese, but none of the adults could solve any item. The reliability coefficient is .76 for the English items. There were insufficient data to calculate accurate reliabilities for the tests developed in the other three languages.

Nonword Decoding Task. Nonwords were used to assess decoding to control for differences in familiarity with real words. A character named Tony was introduced, and children were told that Tony came from a planet called "Starry." It was explained that Tony had to say some secret words in order to start up the spaceship but he forgot how to pronounce the secret codes. Children were asked to help Tony start his spaceship by pronouncing the words for him.

The English items were adapted from the list used by Treiman, Goswami, and Bruck (1990) and incorporated both friendly and unfriendly nonwords. Friendly words have many orthographic neighbors, whereas unfriendly words are more unique, sharing orthographic patterns with relatively few words (Laxon, Coltheart, & Keating, 1988). There were 20 words consisting of 10 friendly and 10 unfriendly words. The friendly words were *gog, fiss, foop, vag, meep, jick, lat, kan, sug, fip*; the unfriendly words were *vepp, lem, fod, paf, veeg, leck, hud, fep, meb, hoog.* Similar lists were constructed in the other languages. In Chinese, the friendly nonwords contained a pronounceable radical in the left side only and the unfriendly nonwords were pronounceable from both the left and right radicals. Ten items were used as Chinese nonwords: Half of the items were friendly, and the other half were unfriendly. The reliability coefficient for this task in English is .77. Reliability for the other three languages could not be computed.

RESULTS

Comparing Group Performance for English Tasks

The mean scores for the English tests are reported in Table 1. The language groups differed significantly in age, $F(3, 128) = 6.09$, $\eta^2 = 0.12$, $MSE = 14.36$, $p < .001$, and post hoc Bonferroni comparisons showed that the Chinese–English bilinguals were younger than the Spanish–English and monolingual groups, Fs > 2.68, ps < .05, with the Hebrew–English group not differing significantly from any of the others. For this reason, all the correlation analyses were conducted by controlling for age. The groups also differed in mean scores on the PPVT–R, $F(3, 128) = 18.00$, $\eta^2 = 0.30$, $MSE = 304.47$, $p < .0001$. Bonferroni comparisons revealed that the monolinguals and Hebrew–English bilinguals obtained higher scores than the Spanish–English and Chinese–English bilinguals, with no differences within each of these pairs (all significant Fs > 2.68, ps < .05). The groups differed on Forward Digit Span, $F(3, 130) = 6.17$, $\eta^2 = 0.12$, $MSE = 2.41$, $p < .001$, and Bonferroni comparisons showed that the Spanish–English bilinguals and the monolinguals scored higher than the Hebrew–English bilinguals and the Chinese–English bilinguals,

TABLE 1
Mean Scores and Standard Deviations on Background Measures and English Tasks

	Monolingual		*Bilingual*					
	English		*Chinese–English*		*Hebrew–English*		*Spanish–English*	
Variables	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Sample size	40		29		30		33	
Age in months	81.13	3.86	78.66	4.38	80.80	3.79	82.76	3.07
PPVT–R standard scores (English)	112.60	15.96	93.97	16.83	102.47	16.96	83.48	19.98
Forward digit span	8.23	1.56	7.45	1.70	7.07	1.51	8.55	1.44
Backward digit span	3.36	1.40	3.31	1.42	3.57	1.10	3.55	1.06
Phoneme counting	5.60	2.00	5.45	1.99	8.57	1.17	7.45	1.18
Nonword decoding	12.18	4.65	12.00	6.39	16.40	3.40	14.64	3.81

Note. PPVT–R = Peabody Picture Vocabulary Test–Revised.

again with no difference within each pair (all significant *F*s > 2.68, *p*s < .05). There was a group effect in the phoneme counting task, $F(3, 131) = 26.45$, $\eta^2 = 0.38$, $MSE = 2.74$, $p < .0001$, because the Hebrew–English and Spanish–English bilinguals outperformed the Chinese–English bilinguals and the English monolinguals, with no differences within pairs (all significant *F*s > 2.68, *p*s < .05). There was no significant difference between groups on the Backward Digit Span, $F < 1$.

The mean scores for nonword decoding are also shown in Table 1. The initial analysis was a multivariate analysis of variance using friendly and unfriendly words as two dependent variables across groups. The analysis showed that the friendly and unfriendly words were strongly correlated, $r(124) = .74$, and that the effect of group was the same for both with no interactions. Therefore, the two types of words were combined into a single variable and examined in a one-way analysis of variance that showed an effect of group, $F(3, 128) = 6.95$, $\eta^2 = 0.14$, $MSE = 21.75$, $p < .001$. A post hoc Bonferroni comparison indicated that the Hebrew bilinguals scored higher than all the other groups and the English monolinguals and Chinese bilinguals scored the lowest, with the Spanish bilinguals not differing from either of these (all significant *F*s > 2.68, *p*s < .05).

Because the children in the four groups differed in their initial abilities that relate to decoding success, an analysis of covariance (ANCOVA) was conducted to compare decoding scores across groups, controlling for the effects of age, phoneme counting, Forward Digit Span, and PPVT–R standard scores. The intention was to isolate the effect of bilingualism on decoding by statistically equating groups on variables that relate to decoding ability and leave bilingualism as the only identifiable difference among the groups. The overall ANCOVA produced a significant language group effect, using an $\alpha = .05$, $F(3, 124) = 7.64$, $p < .01$, partial $\eta^2 = 0.16$, whereas $\omega^2 = 0.11$ with a 95% confidence interval (CI; Smithson, 2003) of (.04, 0.25). With reference to $\alpha = .05$, a power analysis showed that this *F* test had an estimated power of 97%. The least square means and standard errors obtained from the ANCOVA are presented in Table 2.

All pairwise comparisons were carried out on the covariate-corrected marginal language group means. Using a Bonferroni adjustment for six two-sided comparisons and a family-wise $\alpha = .05$, the critical significance level is .05/12 = .0004. This led to two significant comparisons: English versus Hebrew and English versus Spanish. As a measure of effect size, each pair of mean differences was divided by the square root of *MSE* (Steiger & Fouladi, 1997), leading to an index analogous to that of Cohen's *d*. Such effect size estimates reveal that five of the six comparisons have absolute values of 0.60 or higher. The 95% CIs (Steiger & Fouladi, 1997) for these effect sizes are provided in Table 2.

Checking the residuals of this ANCOVA model, one observation was identified as having both a high leverage and studentized residual. The ANCOVA model and pairwise comparisons were repeated with this observation omitted from the data. Due to the high similarity of results, this observation was retained in the data.

TABLE 2
Bonferroni Pairwise Contrasts Comparing English Nonword Decoding After Controlling for Age, PPVT, Forward Digit Span, and English Phoneme Counting

Group	*LS Means (SE)*	*Comparison*	*t Value (p)*	*Effect Size*	*95% CI for Effect Size*
English	10.63 (0.77)	Hebrew	–5.21 (< .0001)*	–1.26	(–1.75, –0.76)
		Spanish	–5.50 (< .0001)*	–1.29	(–1.78, –0.80)
		Chinese	–2.46 (.02)	–0.60	(–1.08, –0.12)
Hebrew	15.86 (0.89)	Spanish	–0.14 (.89)	–.03	(–1.14, 0.46)
		Chinese	–2.53 (.01)	–0.66	(–1.17, –0.14)
Spanish	16.00 (0.86)	Chinese	–2.72 (.01)	–0.69	(–1.15, –0.19)
Chinese	13.13 (0.86)	—			

Note. PPVT = Peabody Picture Vocabulary Test; LS = least square; CI = confidence interval.
*$p < .01$.

Comparing Skill Across Languages for Bilinguals

To examine the possibility of transfer of skill across languages, the relationship between nonword decoding scores in the two languages was examined for the bilinguals. The mean scores for the non-English tasks are presented in Table 3. The PPVT scores obtained in the non-English languages were considered as approximations of the bilingual children's levels of receptive vocabulary in the other language. These raw scores for the three bilingual groups were not significantly different from each other, $F(2, 91) = 0.89$, *ns,* reflecting comparable levels of receptive vocabulary in the three groups of bilingual children. Moreover, all the bilingual groups performed better in the English PPVT than in the non-English versions, $t(28) = 9.06$, $t(32) = 15.31$, and $t(29) = 5.18$, all ps < .0001, for Chinese–English, Hebrew–English, and Spanish–English groups respectively. Bilingual children always demonstrate higher vocabulary scores in one language, usually the language of schooling, a gap that does not close until about fifth grade (Oller & Eilers, 2002).

The correlation in performance across languages for the background tasks are reported in Table 4. The entries on the diagonal indicate the correlation between the same tasks in the two languages. There were no correlations between PPVT scores across languages for any group, rs < .33, *ns*. For all groups, each of Forward and Backward Digit Spans were correlated in the two languages, rs > .38, ps < .03, although not correlated to each other (except for the Spanish group, where Forward Digit Span correlated with all other measures), and phoneme counting was correlated for the two groups who completed this task in both languages.

The scatterplots comparing decoding in English and in the other language are presented in Figure 1. The top row shows the raw correlations between decoding

TABLE 3
Mean Scores and Standard Deviations on Non-English Tasks

	Chinese		*Hebrew*		*Spanish*	
Language	*M*	*SD*	*M*	*SD*	*M*	*SD*
PPVT–R raw scores	40.38	15.67	37.53	8.50	42.58	18.50
Forward Digit Span	8.72	1.56	5.90	1.06	7.15	1.15
Backward Digit Span	3.24	1.60	3.40	0.72	3.36	0.96
Phoneme counting	NA	NA	7.97	2.66	8.48	1.48
Phoneme onset deletion	NA	NA	9.93	0.25	8.06	1.74
Nonword decoding	6.31	3.31	16.80	2.50	13.00	4.56

Note. Chinese nonword decoding is out of 10, and Hebrew and Spanish decoding are out of 20. PPVT–R = Peabody Picture Vocabulary Test–Revised.

TABLE 4
Zero-Order Cross-Linguistic Correlations for Background Tasks by Language Group

	Non-English Variables			
English Variables	*1*	*2*	*3*	*4*
Chinese–English bilinguals[a]				
PPVT–R raw	.33	.26	.10	—
Forward Digit Span	.39*	.48**	.20	—
Backward Digit Span	.15	.22	.55**	—
Phoneme counting	–.15	.13	.01	—
HebrewEnglish bilinguals[b]				
PPVT–R raw	.07	.05	.33	–.14
Forward Digit Span	.30	.54**	.29	.29
Backward Digit Span	–.11	.14	.48**	.03
Phoneme counting	.02	.02	–.03	.49**
Spanish–English bilinguals[c]				
PPVT–R raw	.19	.43*	.24	.29
Forward Digit Span	.12	.38*	.30	.27
Backward Digit Span	.03	.54**	.44*	.03
Phoneme counting	.14	.34*	.37*	.63**

[a]$n = 29$. [b]$n = 30$. [c]$n = 33$.
*$p < .05$. **$p < .01$.

scores represented by Pearson correlations of –.09 ($n = 29$, $p = .64$) for the Chinese children, .57 ($n = 30$, $p < .01$) for the Hebrew children, and .72 ($n = 33$, $p < .01$) for the Spanish children. The scatterplots in the second row depict the relation between decoding scores after partialling out age. Within each group, we regressed the English decoding score on age and obtained the residuals, then regressed the

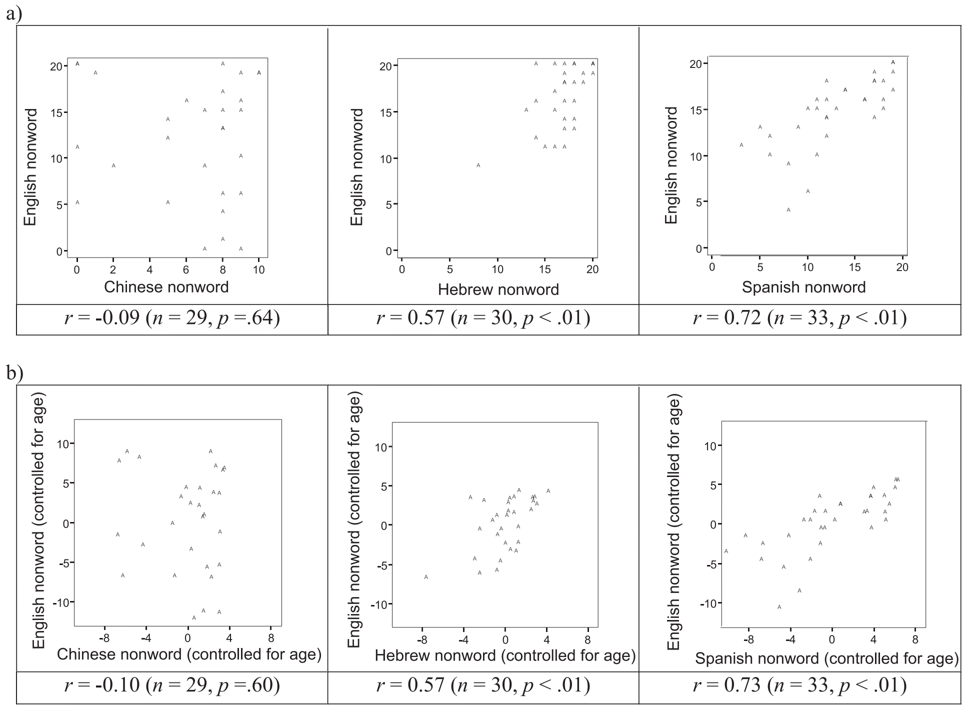

FIGURE 1 Scatterplots between nonword decoding tasks in English and the other languages for the bilingual groups. (a) Zero-ordered correlates between reading scores. (b) First-ordered correlations between reading scores controlling for age in months.

decoding scores of the other language on age and obtained the residuals, the correlation of these two sets of residuals gives us the partial correlation of the decoding scores. The partial correlations are –.10 ($n = 29$, $p = .60$), .57 ($n = 30$, $p < .01$), and .73 ($n = 33$, $p < .01$), for the Chinese, Hebrew, and Spanish children, respectively.

By applying Fisher's z transformation (e.g., see Hays, 1994, pp. 649–650), all pairwise comparisons between the age-adjusted correlations were tested with Z tests. The low correlation for the Chinese group was significantly different from both the Hebrew, $Z = 2.70$, $p < .01$, and Spanish, $Z = 3.82$, $p < .01$, groups, although these latter two were not different from each other, $Z = –1.06$, $p = .29$. By inverting Fisher's transformation, the 95% CIs between each pairwise comparison were found to be –0.86, –0.2, for the Hebrew–Chinese comparison; –0.91, –0.46, for the Spanish–Chinese comparison; and –0.24, 0.66, for the Hebrew–Spanish comparison. There was one observation in the Hebrew group that influenced the value of the partial correlations, so the partial correlations and consequent pairwise comparisons were recalculated with this observation omitted. The results of this analysis were the same, so the observation was retained in the analyses.

DISCUSSION

All of the children in the study were in first grade and learning to read in school, instructed in English, and speaking English at school and with their friends. The children in three of the groups were also learning to read in another language, and for these children, their family life was conducted primarily in this other language. For two of these groups, Hebrew and Spanish bilinguals, progress in learning to read in English was more advanced than for the children in the other groups. This literacy advantage was found even after progress in skills relevant to literacy was statistically controlled. Similarly, the children in these two advanced groups revealed a strong correlation between their nonword decoding skills in the two languages. The Chinese bilingual children demonstrated some advantage relative to monolinguals in the decoding task once initial skill levels were accounted for, but their performance in the phonological awareness task was at the same level as the monolinguals and significantly lower than that of the other two bilingual groups. Therefore, both facilitation for English reading and transfer of reading across languages was found for bilingual children whose two languages shared a writing system.

In the analysis of raw scores, the Chinese–English bilinguals performed the same as the monolinguals, apparently revealing no benefit of their unique language profile. When initial abilities in the relevant background variables were statistically controlled, the Chinese–English children moved higher and fell numerically between the monolingual children and the two bilingual groups, not reliably different from either. Thus, these children were profiting in some manner from their bilingualism, especially considering the reduced English proficiency with which they began literacy instruction. This pattern points to a clear divide between children in three of the groups, all of whom were bilingual but whose non-English language created a different experience in learning to read.

Many factors conspire to predict children's success in acquiring early literacy skills, so isolating the unique role of bilingualism must be undertaken in the context of these other factors. Does bilingualism, in other words, influence the acquisition of early literacy, all else being equal? Previous research, described earlier, has shown that oral vocabulary, phonological awareness, and working memory contribute to children's early progress in reading, and the bilingual children in this study differed from the monolinguals in their performance on some of these skills. The bilingual deficit in vocabulary replicates a frequently reported finding. There is no obvious connection between working memory and bilingualism but because the differences parallel age differences, it may be age that is responsible for that effect in these results. The bilingual advantage in phonological awareness may reflect a positive consequence of bilingualism, an individual difference that allows some children to become bilingual or an outcome of speaking two oral systems that are phonologically related. Moreover, the two groups with the highest phoneme counting scores also obtained the highest decoding scores, pointing again to the interactive nature of these skills. Therefore, the reasons for group differences may not be the same for the various measures, sometimes signaling a positive consequence of bilingualism (phonological awareness), sometimes a negative consequence (vocabulary), and sometimes indicating no relation at all (working memory). This diversity draws attention to the point that bilingualism is not a simple grouping variable with a homogenous effect on performance. Ultimately, the balance between the vocabulary deficit and possible metalinguistic advantage for the Chinese bilinguals resulted in little overall gain in literacy skill, but the possibility of transfer of reading insight for the bilinguals whose two languages were alphabetic resulted in a clear and significant advantage in progress.

We did not systematically investigate SES, but the Spanish–English group lived in a different neighborhood from the other three, and this neighborhood appeared to be less privileged. If there was a difference in SES, it may affect progress in literacy and other academic skills (Duncan & Seymour, 2000). Therefore, the advantages found for this group on several of the experimental measures, including the decoding task, are particularly compelling. Further research should investigate whether a group of Spanish–English bilinguals more equivalent to the other groups in educational privilege would pull further ahead of the Hebrew–English group.

The bilingual children were all receiving literacy instruction in their other languages as well as in English, so it is possible that their advantage simply reflects more intensive literacy training rather than a consequence of bilingualism. The pattern of results, however, is not consistent with the outcomes expected if hours of literacy training were significant. The Hebrew–English bilinguals received Hebrew literacy instruction daily and the Spanish–English and Chinese–English bilinguals received their non-English instruction weekly, so the prediction is that the Hebrew–English bilinguals would be more proficient in the literacy tasks than the Spanish or Chinese groups, who would not differ from each other. This is not what the results showed. The Hebrew and Spanish groups were the same, and both were more advanced than the Chinese group.

Our interpretation of the results is that bilingualism makes two contributions to children's early acquisition of literacy. The first is a general understanding of reading and its basis in a symbolic system of print. This general understanding can be acquired in any writing system and gives children an essential basis for learning how the system works and how the forms can be decoded into meaningful language. To this end, all the bilingual children showed some advantage relative to the monolinguals. The advantage given to the Chinese–English bilinguals was modest, and under conservative criteria for pairwise comparisons, their adjusted scores were no better than those of monolinguals.

The second advantage of bilingualism is the potential for transfer of reading principles across the languages. This transfer is facilitated if the two languages are written in the same system, enabling children to transfer the strategies and expertise that they build up in one of the languages. Thus, the two groups whose two languages were both based on an alphabetic principle enjoyed the additional advantage of applying the concepts of reading that they learn to their two languages, enhancing both and boosting their passage into literacy.

ACKNOWLEDGMENTS

The research reported in this article was funded by Grant A2559 from the Natural Sciences and Engineering Research Council of Canada to Ellen Bialystok. We are grateful for the assistance of Liat Ben Zeev and Carmen Romero, who collected data from two of the bilingual groups.

REFERENCES

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Ben-Zeev, S. (1977). The influence of bilingualism on cognitive strategy and cognitive development. *Child Development, 48*, 1009–1018.

Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Developmental Psychology, 24*, 560–567.

Bialystok, E., & Herman, J. (1999). Does bilingualism matter in early literacy? *Bilingualism: Language and Cognition, 2,* 35–44.

Bialystok, E., Majumder, S., & Martin, M. (2003). Developing phonological awareness: Is there a bilingual advantage? *Applied Psycholinguistics, 24*, 27–44.

Bruck, M., & Genesee, F. (1995). Phonological awareness in young second language learners. *Journal of Child Language, 22,* 307–324.

Bryant, P., & Goswami, U. (1987). Beyond grapheme-phoneme correspondence. *Current Psychology of Cognition, 7*, 439–443.

Campbell, R. & Sais, E. (1995). Accelerated metalinguistic (phonological) awareness in bilingual children. *British Journal of Developmental Psychology, 13*, 61–68.

Cheung, H., Chen, H.-C., Lai, C. Y., Wong, O. C., & Hills, M. (2001). The development of phonological awareness: Effects of spoken language experience and orthography. *Cognition, 81,* 227–241.

Cobo–Lewis, A. B., Pearson, P. Z., Eilers, R. E., & Umbel, V. C. (2002). Effects of bilingualism and bilingual education on oral and written Spanish skills: A multifactor study of standardized test outcomes. In D. K. Oller & R. E. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 98–117). Clevedon, UK: Multilingual Matters.

Comeau, L., Cormier, P., Grandmaison, E., & Lacroix, D. (1999). A longitudinal study of phonological processing skills in children learning to read in a second language. *Journal of Educational Psychology, 91,* 29–43.

Coulmas, F. (1989). *The writing systems of the world*. Oxford, England: Basil Blackwell.

Dale, P. S., Crain-Thoreson, C., & Robinson, N. M. (1995). Linguistic precocity and the development of reading: The role of extralinguistic factors. *Applied Psycholinguistics, 16*, 173–187.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test–Revised.* Circle Pines, MN: American Guidance Center.

Dunn, L. M., Lugo, D. E., Padilla, E. R., & Dunn, L. M. (1986). *Test de Vocabulario en Imagenes Peabody*. Circle Pines, MN: AGS Publishing.

Duncan, L. G., & Seymour, P. H. K. (2000). Socio-economic differences in foundation-level literacy. *British Journal of Psychology, 91,* 145–166.

Durgunoğlu, A. Y. (1988). Repetition, semantic priming and stimulus quality: Implications for the interactive-compensatory reading model. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14*, 597–604.

Durgunoğlu, A. Y., Nagy, W. E., & Hancin-Bhatt, B. J. (1993). Cross-language transfer of phonological awareness. *Journal of Educational Psychology, 85*, 453–465.

Geva, E., & Siegel, L. (2000). Orthographic and cognitive factors in the concurrent development of basic reading skills in two languages. *Reading and Writing, 12*, 1–30.

Geva, E., Wade-Woolley, L., & Shany, M. (1997). Development of reading efficiency in first and second language. *Scientific Studies of Reading, 1,* 119–144.

Gholamain, M., & Geva, E. (1999). Orthographic and cognitive factors in the concurrent development of basic reading skills in English and Persian. *Language Learning, 49*, 183–217.

Gottardo, A., Yan, B., Siegel L. S., & Wade-Woolley, L. (2001). Factors related to English reading performance in children with Chinese as a first language: More evidence of cross- language transfer of phonological processing. *Journal of Educational Psychology, 93,* 530–542.

Hanley, J. R., Tzeng, O., & Huang, H. –S. (1999). Learning to read Chinese. In M. Harris & G. Hatano (Eds.), *Learning to read and write: A cross-linguistic perspective. Cambridge studies in cognitive and perceptual development* (pp. 173–195). New York: Cambridge University Press.

Hays, W. L. (1994). *Statistics* (5th ed.). Belmont, CA: Wadsworth.

Ho, C. S. H., & Bryant, P. (1997). Phonological skills are important in learning to read Chinese. *Developmental Psychology, 33,* 946–951.

Huang, H. S., & Hanley, J. R. (1994). Phonological awareness and visual skills in learning to read Chinese and English. *Cognition, 54*, 73–98.

Laxon, V. J., Colheart, V., & Keating, C (1988). Children find friendly words friendly too: Words with many orthographic neighbours are easier to read and spell. *British Journal of Educational Psychology, 58,* 103–119.

Liberman, I. Y., Shankweiler, D., Liberman, A. M., Fowler, C., & Fischer, F. W. (1977). Phonetic segmentation and recoding in the beginning reader. In A. S. Reber & D. Scarborough (Eds.), *Toward a psychology of reading: The proceedings of the CUNY Conference* (pp. 207–225). Hillsdale: NJ: Lawrence Erlbaum Associates, Inc.

Lindsey, K. A., Manis, F. R., & Bailey, C. E. (2003). Prediction of first-grade reading in Spanish- speaking English-language learners. *Journal of Educational Psychology, 95,* 482–494.

Luk, G. (2003). *Exploring the latent factors behind inter-language correlations in reading and phonological awareness.* Unpublished master's thesis, York University, Toronto, Ontario, Canada.

McBride-Chang, C., & Chang, L. (1995). Memory, print exposure, and metacognition: components of reading in Chinese children. *International Journal of Psychology, 30,* 607–616.

Merriman, W. E., & Kutlesic, V. (1993). Bilingual and monolingual children's use of two lexical acquisition heuristics. *Applied Psycholinguistics, 14,* 229–249.

Morais, J. (1987). Phonetic awareness and reading acquisition. *Psychological Research, 49*, 147– 152.

Oller, D. K., & Eilers, R. E. (Eds.). (2002). *Language and literacy in bilingual children.* Clevedon: Multilingual Matters.

Perfetti, C. A., Beck, I., Bell, L. C., & Hughes, C. (1988). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. In K. E. Stanovich (Ed.), *Children's reading and the development of phonological awareness* (pp. 39–75). Detroit, MI: Wayne State University Press.

Rosenblum, T., & Pinker, S. A. (1983). Word magic revisited: Monolingual and bilingual children's understanding of the word-object relationship. *Child Development, 54,* 773–780.

Shu, H., Anderson, R. C., & Wu, N. (2000). Phonetic awareness: Knowledge of orthography- phonology relationships in the character acquisition of Chinese children. *Journal of Educational Psychology, 92,* 56–62.

Smithson, M. (2003). *Confidence intervals*. Thousand Oaks, CA: Sage.

Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research, 56,* 72–110.

Stanovich, K. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 360–406.

Steiger, J. H., & Fouladi, R. T. (1997). Noncentrality interval estimation and the evaluation of statistical models. In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 221–257). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Treiman, R., Goswami, U., & Bruck, M. (1990). Not all nonwords are alike: Implications for reading development and theory. *Memory and Cognition, 18,* 559–567.

Umbel, V. M., Pearson, B. Z., Fernández, M. C., & Oller, D. K. (1992). Measuring bilingual children's receptive vocabularies. *Child Development, 63,* 1012–1020.

Wade-Woolley, L., & Geva, E. (2000). Processing novel phonemic contrasts in the acquisition of L2 word reading. *Scientific Studies of Reading, 4*, 295–311.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology, 30,* 73–87.

Wechsler, D. (1974). *Wechsler Intelligence Scale for Children–Revised.* New York: Psychological Corporation.

Yelland, G. W., Pollard, J., & Mercuri, A. (1993). The metalinguistic benefits of limited contact with a second language. *Applied Psycholinguistics, 14*, 423–444.

Manuscript received August 18, 2003
Accepted June 14, 2004