Applied Psycholinguistics **21** (2000), 451–471
Printed in the United States of America

Bilingual is as bilingual does: Metalinguistic abilities of Arabic-speaking children

ZOHAR EVIATAR and RAPHIQ IBRAHIM
University of Haifa

ADDRESS FOR CORRESPONDENCE
Zohar Eviatar, Institute of Information Processing and Decision Making, University of Haifa, Mount Carmel, Haifa, Israel 31905. Email: zohare@psy.haifa.ac.il

ABSTRACT
The study explores the effects of the relationship between exposure to two languages in childhood and metalinguistic abilities. Arabic-speaking children who had been exposed to both spoken and literary Arabic were compared to Russian–Hebrew bilinguals and Hebrew monolinguals. All of the children were in kindergarten or first grade. The tests included language arbitrariness, phonological awareness, and vocabulary. As compared to the Hebrew monolinguals, the Russian–Hebrew bilinguals revealed the following pattern: higher performance on arbitrariness and phonological awareness tasks and lower performance on the vocabulary measure. The results of the Arab children mimicked those of the Russian–Hebrew bilinguals and differed from those of the Hebrew monolinguals. We conclude that exposure to literary Arabic requires the same intensive language analyses as those demanded of children exposed to languages as different as Russian and Hebrew.

A significant amount of research has suggested that exposure to more than one language at an early age results in a heightened awareness of the arbitrary, phonological, and grammatical aspects of language (e.g., Bialystok, 1991). These metalinguistic abilities have been tapped by testing children's performance on symbol substitution tasks (e.g., Ben Zeev, 1977; Dash & Mishra, 1992; Edwards & Christophersen, 1988), tasks requiring the manipulation of sublexical phonological elements (e.g., Bruck & Genesee, 1995; Campbell & Sais, 1995; Titone, 1994), and tasks requiring the detection of ambiguity and grammaticality (e.g., Galambos & Goldin-Meadow, 1990; Galambos & Hakuta, 1988). All of these require an awareness of language as a system and the ability to access and manipulate knowledge about that system (Bialystok & Ryan, 1985).

One of the central debates in this area of research concerns the degree of bilingualism necessary for the emergence of metalinguistic advantage. The concept of degree of bilingualism can be seen in two ways. One view that has received attention has defined "degree" as the level of facility in, or exposure to, two languages. Here the results are equivocal: some researchers have re-

ported positive effects of bilingualism with only minimal exposure to a foreign language (e.g., Yelland, Pollard, & Mercuri, 1993), while others have suggested that these effects are only seen in children who have achieved high facility in both languages (Cummins, 1987; Ricciardelli, 1992). However, we are interested in a different definition, one that defines "degree" as an index of the difference between the languages of the bilingual. This view is based on the continuum between dialects of a single language, diglossic situations, and bilingualism and is derived, in part, on the conclusions from more than 30 years of neurolinguistic research, which suggests that the languages of multilinguals are not represented differently in the brain than is the single language of monolinguals. Paradis (1998) suggested that the representation of more than one language differs in a quantitative rather than a qualitative manner from the representation of a single language. The goal of the present research is to investigate how the degree of difference between the linguistic systems children use affects metalinguistic advantage.

The research presented here focuses on the part of the question dealing with the degree of difference between the language systems to which children are exposed. A group of monolingual Hebrew speakers was compared with two groups of bilinguals: children of immigrants from the former Soviet Union whose home language was Russian but who were born in Israel and integrated in Hebrew-speaking schools, and children whose native language was Arabic and who had not been systematically exposed to any other language. Arabic was chosen because of its critical linguistic and cognitive status. Arabic has two forms: literary Arabic (also known as Modern Standard Arabic) is universally used in the Arab world for formal communication and writing; spoken Arabic is a local dialect, which has no written form. The spoken dialect is the native language of all native speakers of Arabic, while literary Arabic is taught in school along with instruction in reading and writing. Although sharing a limited subgroup of words, the two forms of Arabic are phonologically, morphologically, and syntactically different. For example, certain vowels such as /ɛ/ and /o/ exist in spoken Arabic but not literary Arabic. In spoken Arabic, words may begin with two consecutive consonants or with a consonant and a schwa, whereas this is illegal in literary Arabic. The two forms utilize different inflections (such as plural markings) and different insertion rules for function words. They have different word order constraints with regard to sentence structure. As spoken Arabic has no written form, literary Arabic becomes part of everyday life: it is the language in which news is reported (both written and oral) and the language of prayer and of public occasions. In addition, preliterate children are exposed to literary Arabic: children's books are necessarily written in that form, and many cartoon shows and children's television programs utilize literary Arabic to various extents.

A debate exists as to whether the two forms of Arabic represent different languages or whether this is a diglossic situation in which a language community uses two forms of the same language (such as a spoken form and a normative or high form, as in Tamil) (Eid, 1990). Ibrahim and Bentin (2000) addressed this issue directly. They examined the relationship between the two forms of Arabic in adults by comparing semantic priming and repetition effects within

the native language (spoken Arabic). Effects were found when the primes were either in literary Arabic or Hebrew and the targets were in spoken Arabic, and vice versa. Using lexical decisions for auditorally presented words, they found that the semantic priming effect was larger when both the prime and the target were in spoken Arabic than when the prime was in literary Arabic or Hebrew and the targets were in spoken Arabic. Large repetition effects at relatively long lags were found within spoken Arabic, but they were absent when the repetition involved translation equivalents in Hebrew or literary Arabic. These findings suggest that, despite their intensive every day use and psychological proximity, spoken and literary Arabic are represented in two different lexicons in the cognitive system of the native Arabic speaker. However, in line with the close relationship between the two forms of Arabic, as described by Ferguson (1959), the statistical differences found between Hebrew, literary Arabic, and spoken Arabic indicate a larger proximity in the cognitive system between the two forms of Arabic than between Hebrew and spoken Arabic.

Thus, we have a group of bona fide bilinguals, as Russian and Hebrew are different languages by any account, and a group of questionable bilinguals, as the relations between the two forms of Arabic are controversial. In addition, our bilingual groups differed in their levels of facility in the second language. The Russian–Hebrew bilinguals had been integrated in Hebrew-speaking schools and lived in an environment where Hebrew was the majority language. The Arab children had been exposed to literary Arabic to a limited extent (via books and television) but had not been exposed to it in a systematic way in their daily lives. (The day-to-day business of living is conducted in spoken Arabic outside of the home as well as at home.) We posed the following question: given that the Russian–Hebrew bilingual children would reveal different performance patterns from the Hebrew monolingual children (as a result of acquiring two rather than one linguistic systems), would preliterate and newly literate Arab children be more similar to the bilinguals or the monolinguals? That is, would exposure to literary Arabic, in addition to their native spoken Arabic, result in the Arab children developing sensitivity to language arbitrariness as well as phonological detection and manipulation abilities similar to those of the bilingual children? The direction of our question is different from the usual one in the literature. The common question is, how does early exposure to two languages affect performance on tests of metalinguistic abilities? Our approach is from the opposite direction. Given that exposure to two languages results in a specific pattern of performance on three types of tests (arbitrariness, phonological awareness, and vocabulary size) and that this pattern is interpreted as reflecting metalinguistic awareness, would children exposed to the two forms of Arabic show this pattern? That is, are the two forms of Arabic different enough from each other to result in the pattern typical of bilingualism?

LANGUAGE TESTS

We tested the children on two central metalinguistic abilities – the arbitrary nature of language and phonemic awareness – and on a vocabulary measure. In addition, the parents of the Arab children were asked to fill out a questionnaire assessing the degree of exposure their child had had to literary Arabic.

Language arbitrariness

The arbitrary nature of language is the understanding of word–referent relations, which imposes high demands on the control of attention and the capacity to detect and correct syntactic and semantic violations. For example, in the symbol substitution task created by Ben Zeev (1977), every time the examiner says the word *they*, the subject says the word *spaghetti*. Thus, the child's response to "They are good children" is "Spaghetti are good children." Here we used a task based on the sun/moon task created by Piaget (1929), but we focused on word substitution rather than semantic violations (e.g., "In this game we say 'sun' instead of 'moon'; what do you see in the sky before you go to bed at night?" The correct response is "the sun."). To succeed in this task, the child must grasp the idea that language is arbitrary and subject to change. The child must suppress the normative answer and remember the rule defined by the experimenter. This type of task has been the most consistent in revealing better performance by bilinguals than matched monolinguals (e.g., Edwards & Christophersen, 1988; Titone, 1994).

Phonological awareness

Phonological awareness refers to knowledge about the phonological structure of spoken words and is tested by asking children to detect, isolate, or manipulate subword phonological segments. We used three tests. Two tests examined the ability to isolate phonemes from a spoken word: one required identification of the first sound, and the other required identification of the last sound (e.g., Bentin, Hammer, & Cahan, 1991; Wallach & Wallach, 1976). The third test examined the ability to manipulate sublexical units. In the version for kindergarten, the children were asked to identify the sounds that are left after deletion of either a syllable or all but the final phoneme: for example, "If we say the word MARKET and take off the MARK, what is left?" (Bentin & Leshem, 1993). In the first grade version, the children were given a word and asked to say it, deleting one syllable in either initial, middle, or final position: for example, "Say the word MARKET without ARK" (the answer is MET). All of these tests were originally constructed and validated in Hebrew. Hebrew and Arabic are both Semitic languages and are similar in their morphophonemic structure. It was therefore possible to construct the stimulus lists in Arabic on the basis of the existing tests in Hebrew. The lists were equated on syllable structure and number of sounds. Performance on all of these tests in Hebrew has been found to correlate highly with reading ability and age (Ben Dror & Shany, 1999).

These tasks were chosen for several reasons. The first is that there is a controversy about whether phonological awareness can arise in the early stages of bilingualism or whether it requires high levels of facility (e.g., Cummins, 1987; Ricciardelli, 1992; Yelland et al., 1993). The second is that a large body of research has shown that children's explicit awareness of the phonemic structure of spoken language is related to reading experience (e.g., Goswami & Bryant, 1990). Of particular interest is Campbell and Sais's (1995) finding that bilingual prereaders are better than matched monolinguals at tasks of phonological aware-

ness; they suggested that exposure to another language increases the ability to manipulate sublexical units even before the onset of literacy. On the other hand, Edwards and Christophersen (1988) showed that, although bilingualism is related to advantages on tests of language arbitrariness, it is literacy – not bilingualism – that is related to advantages on tests of phonological awareness. Bruck and Genesee (1995) showed that the bilingual advantage interacts with literacy and age, whereby bilinguals have higher scores on some measures of phonological awareness in kindergarten but not first grade. Testing kindergartners (preliterate) and first graders (newly literate), we looked for the differential effects of literacy, facility, and language experience.

Vocabulary measure

Comparing monolinguals and bilinguals, a number of researchers have found that monolingual children have a larger vocabulary than bilingual children in the dominant language (e.g., Abudarham, 1997; Doyle, Champagne, & Segalowitz, 1978). Having to share their language experiences between two languages, bilinguals have less opportunity for experience with the vocabulary of either, such that they achieve lower scores, compared to monolinguals, in both their first and second languages.

The pattern of the effects of bilingualism can be characterized as follows: bilinguals often have higher scores than monolinguals on tests of language arbitrariness and phonemic awareness and lower scores than monolinguals on tests of vocabulary size. This is the pattern we expected to find between the Hebrew monolinguals and the Russian–Hebrew bilinguals. Our goal was to use this pattern as a marker to assess the status of literary Arabic in the cognitive system of Arab children.

METHOD

Participants

The participants were 116 children sampled from three populations in the northern region of Israel. All of the participants lived in villages. The Hebrew monolinguals and the Russian–Hebrew bilinguals lived in the same large village (population = 10,000) and attended the village public schools. The Arab children lived nearby (all within a 10 km radius of the Jewish village) in four separate villages (average population of each village = 3,000) and attended the village public schools. Although we did not collect socioeconomic data on our participants specifically, the overall socioeconomic status of the Jewish and Arab villages from which we sampled was similar. We tested 40 native Arabic speakers, 40 monolingual Hebrew speakers, and 36 bilingual Russian–Hebrew speakers. Half of the children from each population were in kindergarten, and half were in first grade. Only children between 4;10 and 5;3 were included in the kindergarten group; children between 6;10 and 7;3 were included in the first grade group.

None of the children attended a mixed school (with both Arabic- and Hebrew-

speaking populations) or lived in a mixed neighborhood. This was done to minimize the exposure of the Arab children to Hebrew, which is the majority language of the country (Arab schools begin teaching Hebrew in second grade). In the majority of this rural Arab population, exposure to media in Hebrew is minimal, as newspapers and TV programs in Arabic are available and preferred.

Both the monolingual and bilingual Hebrew speakers were being taught in Hebrew and were tested in Hebrew. The Arab children were being taught in Arabic and were tested in spoken Arabic (their native language). None of the children suffered from known neurological, emotional, or attention disorders. Only children without known reading disabilities were selected for the first grade group. All the groups consisted of a similar proportion of males and females.

MATERIALS

The Arab children were tested in spoken Arabic (both stimuli and instructions), and the Hebrew monolinguals and Russian–Hebrew bilinguals were tested in Hebrew.

Arbitrariness test

The child was given the following instructions. "We are going to play a game in which we exchange one word for another. I will ask you questions and you answer with the exchanged words. For example, we'll call the sun the moon and the moon the sun. And now the question: when you go to sleep at night, what do you see in the sky? The answer is _____." There were 10 items in this test. Translations of these items are included in Appendix 1.

Phonological awareness tests

All of the words in each language were familiar to the speakers of that language. All of the stimuli and instructions in Arabic were in spoken Arabic. Transliterations of the lists of words in Hebrew and Arabic are included in Appendix 2.

Initial phoneme detection. The children were asked to identify the first sound in a word spoken by the experimenter: "What is the first sound in the word _____?" The test included 20 words: for example, Hebrew: SEEKA 'pin'/ Arabic: LAHEM 'meat'. Syllables were not accepted as correct: for example, the answer SEE to SEEKA would be counted as incorrect; the answer SUH or S would be scored as correct.

Final phoneme detection. The children were asked to identify the last sound in a word spoken by the experimenter: "What is the last sound in the word _____?" The test included 20 words: for example, Hebrew: MEETA 'bed'/Arabic: HILU 'sweet'. Syllables were not accepted as correct: for example, the answer TA to MEETA would be counted as incorrect; the answer AH would be scored as correct.

Phoneme/syllable deletion. This test had two versions, one for kindergartners and one for first graders. In the kindergarten version, the experimenter said a word and then repeated it, omitting the last phoneme or syllable. The task of the child was to identify the omitted sound. In the first grade version, the children were presented with a word and then asked to generate the word, deleting a phoneme or syllable from either the beginning, middle, or end of the word. In this version, the answer was always a real word. There were 20 items in each version of the task. For example:

(Kindergarten version)
Hebrew: If we say the word MASTIK ['chewing gum'] and take off the MASTI, what is left?
Arabic: If we say the word BATIKH ['watermelon'] and take off the BATI, what is left?
(First grade version)
Hebrew: Say the word MATANA ['present'] without TA. [the answer is MANA 'portion']
Arabic: Say the word SEKEENE ['knife'] without KEE. [the answer is SENE 'year']

Vocabulary test

We used the word definition test from translated versions of the WISC-R into Hebrew or Arabic. The child was presented with a word and was asked to explain what it meant in his/her own words. We used the raw scores, where easy items received scores of either 0 or 1 and more difficult items received scores between 0 and 2. The maximum score was 36. There were 22 items in the test.

Exposure to literary Arabic

The parents of the Arab children filled out a 6-item questionnaire designed to assess the degree of exposure their children had had to this language. The responses were rated on a Likert scale, ranging from (1) minimal exposure to literary Arabic to (5) maximum exposure. Translations of the questions are included in Appendix 3.

PROCEDURE

All of the children were tested individually in a relatively quiet room at school. The session was 30 minutes for the kindergartners and 40 minutes for the first graders (who were given additional reading tests as part of a different study). In each session, the tests were given in a fixed order: final phoneme identification, initial phoneme identification, phoneme/syllable deletion, vocabulary test, and arbitrariness. Each test was preceded by practice trials to verify that the child understood the task. During the practice trials the children were given feedback, and when necessary the task was explained again and additional ex-

amples were given. No feedback was given during the experimental trials. All of the sessions were tape-recorded for later transcription and coding. The Arab children and the Hebrew monolinguals performed all of the tests in their native language (spoken Arabic and Hebrew, respectively). The Russian–Hebrew bilinguals performed all of the tests in their second language (Hebrew).

RESULTS

The performance of the children on the various measures of metalinguistic awareness and vocabulary are reported separately.

Arbitrariness

The number of correct responses on the arbitrariness test was analyzed using a 2×3 between-subject ANOVA. The independent variables were age (kindergarten vs. first grade) and language experience (Arab children vs. Hebrew monolinguals vs. Russian–Hebrew bilinguals). The analysis revealed a significant effect of age, $F(1, 110) = 18.93$, $p < .0001$, with first graders ($M = 8.59$) achieving higher scores than kindergartners ($M = 7.02$). There was also a main effect of language experience, $F(2, 110) = 12.4$, $p < .0001$, with Russian–Hebrew bilinguals achieving the highest score ($M = 8.83$), Hebrew monolinguals achieving the lowest score ($M = 6.65$), and Arab children falling in between ($M = 8.08$). Planned comparisons revealed that the scores of Arab children and Russian–Hebrew bilinguals did not differ from each other ($p > .11$), and that the scores of both groups differed significantly from the scores of Hebrew monolinguals: Arab children versus Hebrew monolinguals, $F(1, 113) = 9.35$, $p < .005$; Russian–Hebrew bilinguals versus Hebrew monolinguals, $F(1, 113) = 20.79$, $p < .0001$. The interaction of age and language experience was not significant ($p > .9$). The data are illustrated in Figure 1.

Phonology

The three measures of phonological awareness were analyzed separately.

Initial phoneme detection. The number of correct answers in the initial phoneme detection task was analyzed with a 2×3 ANOVA using age and language experience as between-group variables. This analysis revealed a main effect of age, $F(1, 110) = 7.22$, $p < .01$, with first graders achieving higher scores than kindergartners (9.88 vs. 7.77), and a main effect of language experience, $F(2, 110) = 63.53$, $p < .0001$, with Arab children achieving the highest score (14.60), Hebrew monolinguals achieving the lowest score (4.10), and Russian–Hebrew bilinguals scoring in between (7.72). The interaction of age and language experience was just significant, $F(2, 110) = 3.06$, $p = .051$ (see Figure 1). Planned comparisons revealed that Arab children had significantly higher scores than the other groups in both age groups, and that Russian–Hebrew bilinguals showed significantly higher performance than Hebrew monolinguals in kindergarten, $F(1, 54) = 16.76$, $p < .0001$, but not first grade ($p > .11$).

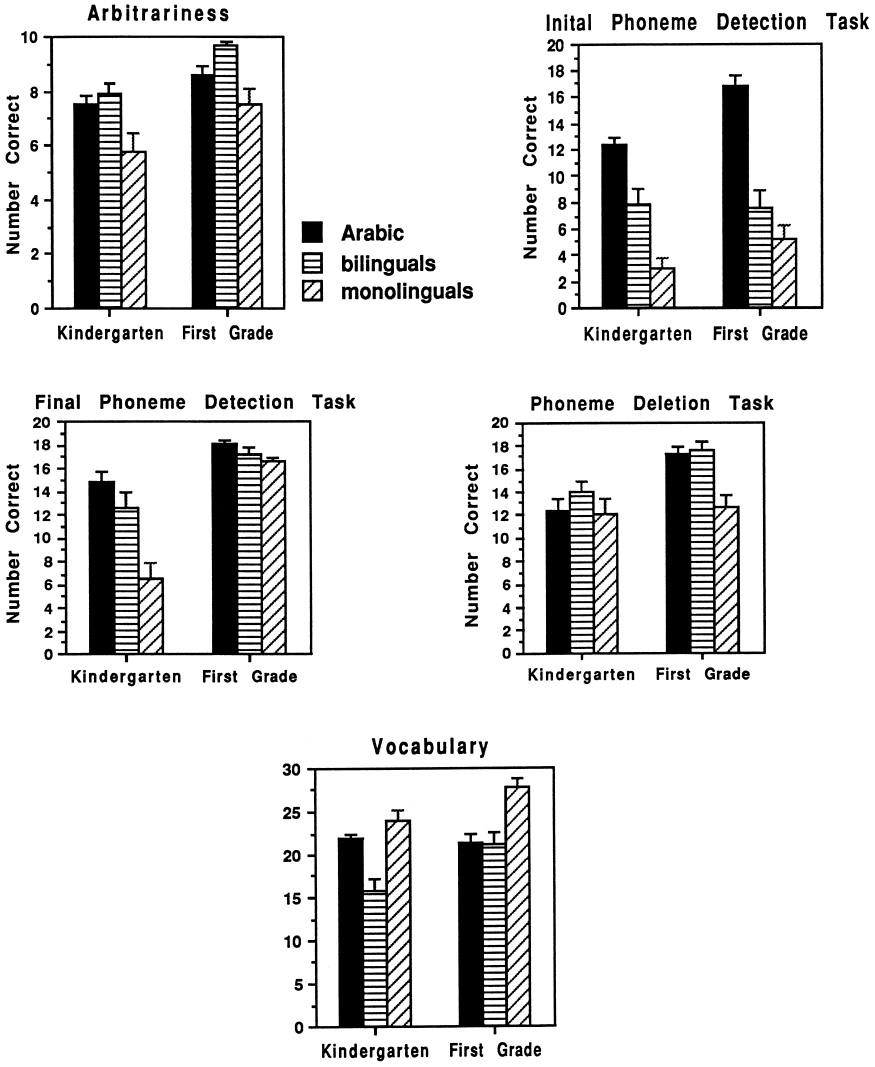

Figure 1. Effects of language experience and age on tests of metalinguistic awareness and vocabulary.

Final phoneme detection. The number of correct responses in the final phoneme detection task was analyzed in the same manner. This analysis revealed a main effect of age, $F(1, 110) = 72.03$, $p < .0001$ (kindergarten = 11.26, first grade = 17.31), a main effect of language experience, $F(2, 110) = 17.19$, $p < .0001$ (Hebrew monolinguals = 11.58, Russian–Hebrew bilinguals = 15.03, Arab children = 16.48), and a significant interaction, $F(2, 110) = 8.85$, $p < .0005$ (see Figure 1). It can be seen that language experience had large effects in kindergarten, whereby both Arab children and Russian–Hebrew bilinguals achieved high scores that did not differ from each other ($p > .19$), but were significantly higher than the scores of Hebrew monolinguals: Arab children versus Hebrew monolinguals, $F(1, 54) = 23.38$, $p < .0001$; Russian–Hebrew bilinguals versus Hebrew monolinguals, $F(1, 54) = 13.14$, $p < .001$. In first grade, only the comparison of Arab children and Hebrew monolinguals was significant, $F(1, 56) = 6.23$, $p < .05$.

Phoneme/syllable deletion task. The results of the phoneme/syllable deletion task are also illustrated in Figure 1. These data were analyzed separately for the two age groups, as they completed different tasks. A one-way analyses of the two age groups revealed different patterns. In kindergarten, language experience had no effect ($p > .7$), but in first grade, language experience had a significant main effect, $F(2, 56) = 12.94$, $p < .0001$. Planned comparisons revealed that Arab children and Russian–Hebrew bilinguals did not differ from each other ($p > .79$), and both groups performed significantly better than Hebrew monolinguals: Arab children versus Hebrew monolinguals, $F(1, 56) = 18.43$, $p < .0001$; Russian–Hebrew bilinguals versus Hebrew monolinguals, $F(1, 56) = 20.19$, $p < .0001$.

Vocabulary

The results of the vocabulary test were analyzed using age and language experience as between-group measures. The mean scores of the groups are illustrated in Figure 1. The analysis revealed a main effect of age, $F(1, 110) = 8.56$, $p < .005$ (kindergarten = 20.81, first grade = 23.47), and a main effect of language experience, $F(2, 110) = 20.48$, $p < .0001$ (Arab children = 21.65, Russian–Hebrew bilinguals = 18.66, Hebrew monolinguals = 25.82). The interaction of age and language experience was also significant, $F(2, 110) = 4.03$, $p < .05$. As can be seen in Figure 1, in kindergarten the vocabulary scores of Arab children were similar to those of Hebrew monolinguals ($p > .19$) and different from those of Russian–Hebrew bilinguals, $F(1, 54) = 14.52$, $p < .0005$, but in first grade their scores were similar to those of Russian–Hebrew bilinguals ($p > .8$) and different from those of Hebrew monolinguals, $F(1, 54) = 15.98$, $p < .0005$. Both Russian–Hebrew bilinguals and Hebrew monolinguals revealed a significant improvement in vocabulary scores in first grade as compared to kindergarten: bilinguals, 15.88 versus 21.16, $F(1, 34) = 7.55$, $p < .05$; monolinguals, 23.9 versus 27.75, $F(1, 38) = 5.59$, $p < .05$. Arab children showed no improvement in first grade compared to kindergarten (21.9 vs. 21.4, $p > .67$).

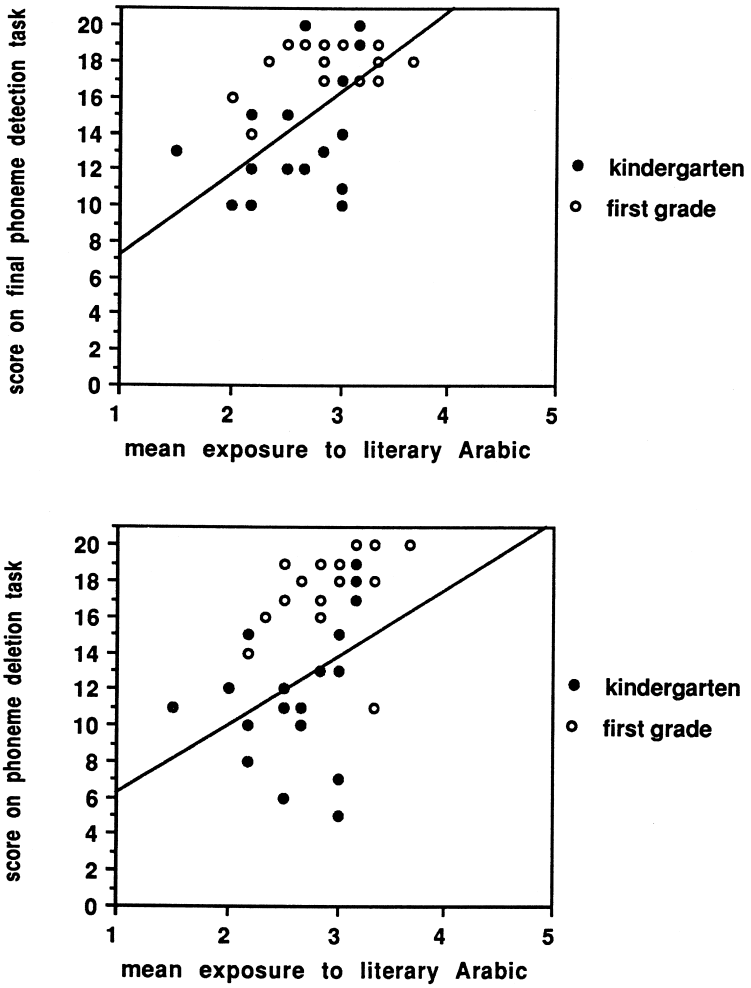

Figure 2. Correlations between the mean score on the language exposure questionnaire and the scores of the Arab children on the final phoneme detection task (top panel) and the phoneme/syllable deletion task (bottom panel).

Metalinguistic measures and exposure to literary Arabic

The mean scores over the six questions in the exposure questionnaire were entered into a correlation analysis with each of the measures of metalinguistic awareness for the 40 Arab children. These analyses revealed a significant positive relationship between exposure to literary Arabic and two of the phonological awareness measures: final phoneme detection, $r(38) = 0.52$, $p < .001$, and phoneme/syllable deletion, $r(38) = 0.47$, $p < .005$. These relationships are illustrated in Figure 2. Analyses of each age group separately revealed that, in the

final phoneme detection task, the relationship between mean exposure score and test score was significant for kindergartners, $r(18) = 0.55$, $p < .05$, and approached significance for first graders, $r(18) = 0.40$, $p = .08$. In the phoneme/syllable deletion task, the relationship between mean exposure and test scores approached significance in kindergarten, $r(18) = 0.42$, $p = .06$, and was significant in first grade, $r(18) = 0.46$, $p < .05$. Comparisons of the correlation coefficients in the two age groups revealed that they did not differ from each other ($p > .05$).

DISCUSSION

The results of the metalinguistic skills and vocabulary measures suggest that preliterate and literate Arab children function as bilinguals as a result of having to deal with the two forms of Arabic. The Russian–Hebrew bilinguals showed the expected pattern resulting from exposure to two languages: higher performance levels in the metalinguistic tests and lower performance levels in the vocabulary measure, as compared to the Hebrew monolinguals. The Arab children's performance levels mimicked those of the Russian–Hebrew bilinguals for the most part, suggesting that exposure to literary Arabic may require the same intensive language analyses as are done for children who are exposed to two languages as different as Hebrew and Russian. This interpretation is supported by the correlation between two of the phonological awareness tests and the estimated degree of exposure to literary Arabic: children whose parents reported higher levels of exposure tended to achieve higher scores on these tests. The results also support Yelland et al.'s (1993) suggestion that even low levels of ability in the second language are related to metalinguistic advantages.

The effects of age are apparent in the measures of phonological awareness: first graders performed at a higher level than kindergartners in all of the measures, presumably as a result of becoming literate. In the initial phoneme detection task, the Russian–Hebrew bilinguals performed better than the Hebrew monolinguals in kindergarten but not first grade because of improvement in the monolinguals' scores, presumably the result of becoming literate (see Figure 1). The advantage of the Arab children in this task was larger in first grade than kindergarten. We believe that this is an artifact of a difference between spoken and literary Arabic: many words in spoken Arabic begin with a schwa (a neutral vowel), whereas this is illegal in literary Arabic. This difference is explicitly taught in kindergarten and emphasized even more in first grade. The generally low level of performance on the Hebrew version of the task (given to the Hebrew monolinguals and the Russian–Hebrew bilinguals) is interpreted as reflecting the manner in which the Hebrew orthography represents phonology. In Hebrew (and Arabic), all verbs and most nouns are written primarily as consonantal roots, which are differently affixed and voweled to form the words of the lexicon (Berman, 1978). Most written materials do not include vowels, although there are four letters in the two languages which, in addition to signifying specific consonants, also specify long vowels; these are called "matres lectionis." However, in some cases it is difficult for the reader to determine whether a dual-function letter represents a vowel or a consonant. When vowels do appear

(in poetry, children's books, and liturgical texts), they are signified by diacritical marks placed above, below, or within the body of the word. Inclusion of these marks specifies the phonological form of the orthographic string, making it transparent in terms of orthography–phonology relations. All of the children were learning to read voweled orthographies (both Hebrew and Arabic). The combination of a consonant and vowel symbol above or below a word results in a specific form of word segmentation, which is evident in very skilled readers of Hebrew (see Ben Dror, Frost, & Bentin, 1995, for a discussion). Thus, most of the children tested in Hebrew responded with a CV when asked to identify the initial phoneme of a word (as did the skilled readers tested by Ben Dror et al.). We believe that the bias to respond in this way in Arabic is overcome by the emphasis put on the difference between spoken and literary Arabic in the initial phoneme.

The advantage of the Arab children and the Russian–Hebrew bilinguals in the final phoneme detection task disappeared in first grade because the Hebrew monolinguals' developing phonological skills resulted in much higher performance levels than in kindergarten (see Figure 1). Interestingly, in the phoneme/syllable deletion task, the effects of age and literacy were in the opposite direction: there was no effect of language experience in kindergarten, but there was a large effect in first grade. However, recall that this test had different versions for kindergarten and first grade. A possible interpretation of this pattern may involve an interaction between bilingualism and task difficulty: perhaps bilingualism relates to better performance on the task demanding higher levels of control processes (the first grade version of the deletion task), as defined by Bialystok (1986), but not on the less difficult task (the kindergarten version of the deletion task). Among the Hebrew speakers, we noted a convergence with the results reported by Bruck and Genesee (1995); in both of the phoneme detection tasks, bilinguals performed better than monolinguals in kindergarten but not first grade, whereas in the phoneme/syllable deletion task, which required the manipulation of syllables, the bilingual advantage was evident in first grade but not kindergarten.

The interaction of age and language experience in the vocabulary measure (see Figure 1) raises an interesting issue. Recall that the Arab children and the Hebrew monolinguals were tested in their respective native languages, whereas the Russian–Hebrew bilinguals were tested in their second language. In kindergarten, the Hebrew monolinguals and the Arab children performed equally well but significantly better than the Russian–Hebrew bilinguals. In first grade, the Hebrew monolinguals and the Russian–Hebrew bilinguals showed significant improvement, but the Arab children did not. Notice that in first grade the improvement in the Russian–Hebrew bilinguals' second language vocabulary scores raised them to the level of the native language scores of the Arab children, and that both groups showed lower performance levels than the Hebrew monolinguals. These results speak to recent discussions in the literature about the validity of comparing bilingual vocabulary in either the first or the second language to monolingual norms, as these tend to underrepresent the vocabulary knowledge of the bilinguals in both languages (e.g., Abudarham, 1997; Umbel, Pearson, Fernandez, & Oller, 1992). Our data support this hypothesis, because,

as compared to the Hebrew monolinguals, both the Arab children and the Russian–Hebrew bilinguals revealed a disadvantage in the vocabulary measure, the former in their native language and the latter in their second language.

Interestingly, although the Arab children revealed higher scores than the Hebrew monolinguals in the test of language arbitrariness, this test was not related to degree of exposure to literary Arabic. We believe that this is a reflection of a qualitative difference between these measures of metalinguistic awareness: phonological awareness, which arises out of specific processes that deal with the structural aspects of a particular language, versus arbitrariness, which is the ability to distinguish between an object or phenomenon in the world and the lexical item created to represent it in a specific language. These are two different forms of knowledge, and we believe they differentially reflect the effects of bilingualism at an early age. Phonological awareness is enhanced by becoming bilingual, but is not necessarily tied to bilingualism. That is, as shown by the many findings that have revealed a relationship between literacy and phonological awareness, it is a result of intensive structural analyses of the sound structure of speech. Learning to read demands an awareness of the sounds of speech, independently of other language experiences. Arbitrariness, on the other hand, may be a specific effect of early bilingualism; learning that there can be more than one form for the same object emphasizes this aspect of languages. Yelland et al. (1993) demonstrated that even exposure to a second language for an hour a week for a few months raises awareness of the arbitrary nature of language.

To summarize, the experiment reported here was designed to explore the effect of the relationship between a bilingual's languages and the emergence of metalinguistic skills in childhood. We choose to study native speakers of Arabic because of the special nature of linguistic and cognitive processes undertaken by literate speakers. Previous results reported by Ibrahim and Bentin (2000) suggested that spoken and literary Arabic have the status of two separate languages in the cognitive systems of adults. The research reported here complements that finding by showing that exposure to literary Arabic in early childhood affects metalinguistic skills in the same manner as that reported for children exposed to two languages.

APPENDIX 1

TEST OF LANGUAGE ARBITRARINESS

Instructions: Now we are going to play a game where we switch one word for another. I will ask you questions, and you will answer after you have switched the words. For example, now we will call the sun the moon and the moon the sun. And now I ask: when you go to sleep at night, what do you see in the sky? The answer is _____.

1. Now we will call a ship a plane and a plane a ship.
   What flies in the air? _____
2. Now we will call a cat a mouse and a mouse a cat.
   Who chases whom? _____
3. Now we will call the bus station school and school a bus station.
   Where do you go every morning? _____
4. Now we will call clean dirty and dirty clean.
   After I fell in the mud my clothes became _____
5. Now we will call a car a frog and a frog a car.
   I looked both ways and then crossed the street because I didn't see a single _____
6. Now we will call a tree a page and a page a tree.
   On Arbor Day, every student plants a _____
7. Now we will call a girl a boy and a boy a girl.
   The parents bought a birthday dress for the little _____
8. Now we will call a hat a towel and a towel a hat.
   When I go for a hike in the sun I wear a _____
9. Now we will call white black and black white.
   This year it snowed by our house and the ground turned _____
10. Now we will call winter summer and summer winter.
    Kids like to go swimming in the _____

APPENDIX 2

TRANSLITERATIONS AND TRANSLATIONS OF THE PHONOLOGICAL TESTS LISTS

Initial phoneme detection

Hebrew	Gloss	Arabic	Gloss
see'ka	pin	sa'ma'ke	fish
o'rez	rice	a'kal	eat
pa'tu'akh	open	fa'tah	opened
tse'va	color	ka'reem	generous
ka'khol	blue	mad'ra'se	school
mad're'ga	stair	os'baa	finger
a'mar	said	la'hem	meat
le'ket	gather	ba'sal	onion
ba'tsal	onion	arth	ground
oo'lay	maybe	e'sem	name
e'retz	land	ja'las	sit
ga'shoom	rainy	da'e'ra	circle
dakh'leel	scarecrow	far'eek	team
vee'lon	curtain	zah'ra	flower
zim'ra	singing	tam'reen	exercise
tar'geel	trick	im'ra'a	woman
ee'sha	woman	re'eh	wind
no'tza	feather	no'kta	point
ree'shon	first	sha'mal	north
sha'lom	hello	ka'mar	moon

Final phoneme detection

Hebrew	Gloss	Arabic	Gloss
ra'ash	noise	ra'mash	blink
gav	back	raf	shelf
kha'veet	barrel	ba'seet	simple
sof	end	shoof	look
ar'gaz	box	gaz	gaz
me'lon	melon	aa'mood	pole
kha'mood	cute	se'fer	zero
tso'far	siren	ha'ram	pity
ma'rom	sky	taj	crown
khag	holiday	bas	kiss
kos	cup	ja'bal	mountain
ga'mal	camel	cha'mees	thursday
kha'mootz	sour	ba'reed	post
ba'ree	healthy	sa'teh	roof
sha'te'each	rug	ma'lek	king
de'vek	glue	he'lou	nice
mee'ta	bed	ka'kao	kakao
ta'leh	lamb	ba'ka'ra	cow
ka'kao	cocoa	sook	market
pa'rah	cow	ba'tal	hero

Deletion tasks

Instructions: I will say a word, then I will take away part of it, and you will tell me what is left: for example, if I say GADOL and take away GADO, what is left?

Kindergarten version

Hebrew		Arabic	
Say	Take away	Say	Take away
mas'tik	masti	ba'teech	batee
ka'peet	kapee	se'keen	sekee
khesh'bon	kheshbo	tho'ban	thoba
pa'nas	pana	le'bas	leba
o'rez	ore	a'mal	ama
yeled	yele	wa'lad	wala
rosh	ro	sam	sa
par'tzuf	partzu	maj'noon	majnoo
san'dal	sanda	ha'deed	hadee
sha'lom	shalo	sa'lam	sala
mad're'ga	madre	mad'ra'se	madra
ba'lon	bal	ba'lon	bal
kee'shut	keesh	ta'boot	tab
pa'kheet	pakh	ja'meel	jam
kvee'sa	kvee	he'san	he
tzee'nor	tzeen	ma'dee'na	madee
ma'kha'ne	makha	sha'dee'da	shadee
savta	sav	sa'boon	sa
kha'vee'la	khavee	ka'see'ra	kasee
ka'tom	kat	ka'lam	kal

First grade version

Hebrew		Arabic	
Say	Take away	Say	Take away
ge'shem	ge	na'shar	na
ma'ta'na	ta	ke'tar	ke
ba'yit	yi	ba'kar	kar
kaf'tor	kaf	sha'mea'	ea'
kha'tee'ma	tee	ja'maa	ja
ya'kar	ya	sa'mak	k
kha'vee'la	vee	ha'kam	ka
ge'zer	ge	ja'mal	ja
ga'nav	na	ra'kas	ra
ka'ba'la	ba	da'rab	da
khad'ron	ron	dah'raj	dah
she'men	me	mak'aad	mak
mad'khom	mad	ya'bes	ya
bad'ran	ran	do'kan	do
mis'par	mis	sa'fa'ra	sa
saf'sal	saf	ba'see'ta	see
kal'mar	mar	se'kee'ne	kee
sha'khor	sha	maj'rooh	maj
shab'lool	shab	ma'ta'ra	ta
ga'mal	ma	mas'naa	naa

APPENDIX 3

QUESTIONNAIRE: EXPOSURE TO LITERARY ARABIC

1. What language do you use in speaking to your child?

1	2	3	4	5
Spoken Arabic only	Spoken Arabic mostly	Spoken and literary equally	Mostly literary Arabic	Literary only

2. To what degree do you read stories to your child versus tell them the story in spoken Arabic?

1	2	3	4	5
Spoken Arabic only	Spoken Arabic mostly	Spoken and literary equally	Mostly literary Arabic	Literary only

3. To what degree do you use translation to spoken Arabic when reading a story?

1	2	3	4	5
Translate immediately	Read and then translate	When I feel s/he doesn't understand	When the child requests	Never translate

4. How often does your child watch cartoons in which the characters speak in literary Arabic?

1	2	3	4	5
Once a week	Twice a week	On alternate days	Once a day	Several times a day

5. How often does your child watch children's TV programs in literary Arabic?

1	2	3	4	5
Once a week	Twice a week	On alternate days	Once a day	Several times a day

6. How often does your child insert words in literary Arabic in everyday speech?

1	2	3	4	5
Never	A little	To a moderate degree	Often	Very often

ACKNOWLEDGMENTS

The research reported in this article was supported by a research grant from the ZEIT Foundation Ebelin and Gerd Bucerius and by a postdoctoral fellowship funded by the Science Ministry of the Government of Israel to the second author.

REFERENCES

Abudarham, S. (1997). The tenability of comparing the receptive lexical proficiency of dual language children with standardized monoglot norms. *Educational Studies, 23*, 127–143.

Ben Dror, I., Frost, R., & Bentin, S. (1995). Orthographic representation and phonemic segmentation in skilled readers: A cross-language comparison. *Psychological Science, 6*, 176–181.

Ben Dror, I., & Shany, M. (1999). *The development of phonological processes among skilled readers*. Unpublished manuscript.

Bentin, S., Hammer, R., & Cahan, S. (1991). The effects of aging and first grade schooling on the development of phonological awareness. *Psychological Science, 2*, 271–274.

Bentin, S., & Leshem, H. (1993). On the interaction between phonological awareness and reading acquisition: It's a two way street. *Annals of Dyslexia, 43*, 125–148.

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

Berman, R. A. (1978). *Modern Hebrew structure*. Tel Aviv: University Publishing.

Bialystok, E. (1986). Children's concept of word. *Journal of Psycholinguistic Research, 15*, 13–32.

Bialystok, E. (1991). *Language processing in bilingual children*. Cambridge: Cambridge University Press.

Bialystok, E., & Ryan, E. B. (1985). Toward a definition of metalinguistic skill. *Merrill-Palmer Quarterly, 31*, 229–251.

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

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

Cummins, J. (1987). Bilingualism, language proficiency, and metalinguistic development. In P. Homel & M. Palij (Eds.), *Childhood bilingualism: Aspects of linguistic, cognitive, and social development* (pp. 57–73). Hillsdale, NJ: Erlbaum.

Dash, U. N., & Mishra, H. C. (1992). Bilingualism and metalinguistic development: Evidence from Kond tribal culture. *Psychological Studies, 37*, 81–87.

Doyle, A., Champagne, M., & Segalowitz, N. (1978). Some issues on the assessment of linguistic consequences of early bilingualism. In M. Paradis (Ed.), *Aspects of bilingualism* (pp. 13–20). Colombia, SC: Hornbeam Press.

Edwards, D., & Christophersen, H. (1988). Bilingualism, literacy and metalinguistic awareness in preschool children. *British Journal of Developmental Psychology, 6*, 235–244.

Eid, M. (1990). Arabic linguistics: The current scene. In M. Eid (Ed.), *Perspectives on Arabic linguistics 1* (pp. 3–37). Amsterdam: John Benjamins.

Ferguson, C. A. (1959). Diglossia. In P. P. Giglioli (Ed.), *Language and social context* (pp. 232–251). London: Penguin.

Galambos, S. J., & Goldin-Meadow, S. (1990). The effects of learning two languages on levels of metalinguistic awareness. *Cognition, 34*, 1–56.

Galambos, S. J., & Hakuta, K. (1988). Subject specific and task specific characteristics of metalinguistic awareness in bilingual children. *Applied Psycholinguistics, 9*, 141–162.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read*. Hove: Erlbaum.

Ibrahim, R., & Bentin, S. (2000). *Bilingual organization of spoken and literary Arabic in native Arabic speakers: Two languages or one?* Manuscript submitted for publication.

Paradis, M. (1998). Aphasia in bilinguals: How atypical is it? In P. Coppens & Y. Lebrun (Eds.), *Aphasia in atypical populations* (pp. 35–66). Mahwah, NJ: Erlbaum.

Piaget, J. (1929). *The child's conception of the world*. London: Routledge & Kegan Paul.

Ricciardelli, L. A. (1992). Bilingualism and cognitive development in relation to threshold theory. *Journal of Psycholinguistic Research, 21*, 301–316.

Titone, R. (1994). Bilingual education and the development of metalinguistic abilities: A research project. *International Journal of Psycholinguistics, 10*, 5–14.

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

Wallach, M., & Wallach, L. (1976). *Teaching all children to read*. Chicago: University of Chicago Press.

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