Applied Psycholinguistics, page 1 of 27, 2012
doi:10.1017/S0142716412000525

Vocabulary growth in second language among immigrant school-aged children in Greece

PANAGIOTIS G. SIMOS, GEORGIOS D. SIDERIDIS, ANGELIKI MOUZAKI, and ASPASIA CHATZIDAKI
University of Crete

MARIA TZEVELEKOU
Institute of Language and Speech Processing

Received: February 16, 2011 Accepted for publication: January 19, 2012

ADDRESS FOR CORRESPONDENCE
Panagiotis G. Simos, Department of Psychology, University of Crete, Rethymno, 74100, Crete.
E-mail: akis.simos@gmail.com

ABSTRACT
The goal of the study was to assess differences between native Greek and bilingual, immigrant children of Albanian descent learning Greek as a second language on a receptive vocabulary measure. Vocabulary measures were obtained at five time points, 6 months apart, from 580 children attending Grades 2–4. Individual variability on both initial performance (intercept) and growth rate (slope) was assessed using hierarchical linear modeling, which included linguistic/ethnic group, parental education (as a socioeconomic status [SES] indicator), gender, and a measure of nonverbal cognitive ability as time-invariant predictors of vocabulary growth. Results indicated that linguistic/ethnic group, parental education, and baseline nonverbal cognitive ability were significant predictors of initial vocabulary scores, whereas only linguistic/ethnic group and nonverbal ability accounted for significant variability in vocabulary growth rates. Additional analyses confirmed that linguistic/ethnic group remained a significant predictor of receptive vocabulary knowledge at both the intercept and the slope levels even after controlling for the initial differences between groups on parental education and block design subtest scores.

Academic achievement among minority children has received much attention in an attempt to clarify which variables interact with individual and group characteristics to result in educational disadvantages that contribute to school failure. Although the political and societal factors involved in the education of minority students form a quite complex investigation arena, research has identified several relevant issues. For instance, the variability in academic achievement among different ethnic groups and the consistent pattern of underachievement for certain groups highlight the heterogeneity of the minority student population and the need for diversifying educational practices (Cummins, 1984). Bilingualism and its impact upon children's linguistic and cognitive development has been another target of rigorous research (Baker, 2011; Bialystok, Luk, & Kwan, 2005; Bialystok & Majumder, 1998; Bialystok, Majumder, & Martin, 2003; Bialystok & Shapero, 2005). However, the notion of bilingual child within the educational context has not been easy to define, since there is a lot of variation in the degree of noted proficiency across the two (or more) languages used, the age and degree of language exposure, and the circumstances of language learning (Rhodes, Ochoa, & Ortiz, 2005). The primary goal of the present study was to assess patterns of vocabulary growth in children born to immigrant families from Albania and attending Greek immersion classrooms. Receptive vocabulary was selected as a proxy for word-level language ability. Although this measure may not share the direct relevance of expressive vocabulary to everyday communication ability, it has the advantage of not penalizing (at least to the same extent as expressive vocabulary measures) an immigrant student's likely reduced oral exposure to the second language (L2) at home. Through random selection of immigrant and native Greek children from the same classrooms and hierarchical linear modeling of between-subjects sources of individual variability (parental education, gender, and nonverbal cognitive ability), we attempted to account for factors that may determine differential vocabulary growth between the two ethnic/linguistic groups.

In early studies bilingualism appeared to negatively affect many aspects of children's adjustment and to be a major drawback on language development and academic achievement (Cummins, 1984). As early as 1966 Macnamara pooled the results from 77 published studies to reach the general conclusion that "bilinguals have a weaker grasp of language than monoglots" (p. 31). Several of the studies reviewed in that report focused on vocabulary and most employed cross-sectional designs and concentrated on early language development. Macnamara invoked four key contributing factors to the presumed linguistic lag of bilingual children behind their monolingual peers: cognitive (interference between the two languages and sharing of memory resources), sociocultural (poor cultural assimilation of bilingual children), family–interpersonal (lack of a language model for L2 in the family), and pragmatic (less time available to learn L2 as compared to the first language [L1]). Macnamara also argued that instruction implemented through an ill-developed L2 leads to academic delays. Subsequent research, however, highlighted the inadequacy of the "linguistic lag" hypothesis, suggesting that bilingualism can positively influence both cognitive and linguistic development (Bialystok, 2001; Cummins, 1976, 1978, 1979, 2000). According to the interdependency hypothesis formulated by Cummins, achieving proficiency in L1 could enhance L2 acquisition if the learner is suitably exposed to L2 and motivated for learning (Cummins, 1979, 2000). Furthermore, a firm knowledge foundation for L1 could facilitate L2 acquisition through the already developed cognitive and linguistic skills under specific circumstances (Leseman, 2000).

Many researchers have argued for the primacy of social factors over linguistic factors in explaining the academic progress of bilingual children (e.g., Tucker, 1977). Data show that sociocultural and socioeconomic factors play a key role in determining acquisition of both L1 and L2. It appears that socioeconomically disadvantaged children may acquire L2 skills and knowledge at a slower pace (Hakuta, Butler, & Witt, 2000). Home–school language switch has a beneficial effect on functional bilingualism in middle-class majority-language children, whereas it may lead to inadequate command of both L1 and L2 in minority-language children. Furthermore, SES seems to interact with language education programs: In most cases majority bilingual children follow immersion programs, where L1 appears in the curriculum, while minority bilingual children follow "submersion" programs, where L1 is not taken into account and sociocultural and attitudinal factors usually hinder integration into the class environment (Cummins, 2000; Hamers & Blanc, 2000).

Another delicate issue concerns the relation between bilingualism and nonverbal cognitive ability. Bilingual children may or may not perform differently than comparable monolinguals on nonverbal intelligence tests and on more targeted tests of problem solving ability, such as the Wechsler Intelligence Scale for Children (WISC) block design subtest. However, even if they do, it is not clear whether differences in either direction reflect levels of intellectual capacity or are attributable to bilingualism (for instance, see Bialystok & Craik, 2010; Bialystok & Majumder, 1998).

VOCABULARY DEVELOPMENT AMONG SCHOOL-AGED BILINGUAL CHILDREN

There is an ongoing discussion over whether bilingualism results in a lower rate of vocabulary development for bilingual students than for monolinguals. The majority of existing studies involve Spanish immigrant children attending US schools, with some researchers reporting lower vocabulary scores for bilingual than monolingual children (Bialystok, 1988; Leseman, 2000) while others fail to find differences in growth rates (Goodz, 1994). Conflicting evidence has been added by more recent studies of lexical acquisition of minority students learning English. For instance, in a study examining the profiles of two groups of fifth graders over a 2-year period, children speaking English as a L2 lagged behind their monolingual peers on word meanings and academic vocabulary (Jean & Geva, 2009), challenging Cummins's (1980) proposal for a 5- to 7-year period required by immigrant children to master academic language.

Vocabulary growth assessment should take into account the developmental level of the students, since the interlanguage fossilization factor plays a significant role in L2 acquisition (Selinker, 1972). The findings, for instance, from English-language learners indicate quite rapid growth at lower elementary grades (where students progress from beginning to middle proficiency levels) in contrast to the slower progress that characterizes upper grades (middle to upper levels of language proficiency; Genesee, Lindholm-Leary, Saunders, & Christian, 2005). Evidence collected from Spanish-speaking students from four US states and their English-only classmates indicates considerable differences in vocabulary knowledge (breadth and depth) despite substantial L2 growth among bilingual students (August, Carlo, Dressler, & Snow, 2005; August et al., 1999). Similar findings were reported by the Early Childhood Study of Language and Literacy Development of Spanish-Speaking Children, where longitudinal data from young Spanish/English children from low SES backgrounds were analyzed (Paez & Rinaldi, 2006; Uccelli & Paez, 2007).

Differences in vocabulary development combined with disadvantaged SES may place bilingual students at risk for literacy difficulties and academic underachievement, in light of the increasing recognition of the importance of vocabulary to academic achievement (August & Shanahan, 2006; August et al., 2006; Snow, Burns, & Griffin, 1998; Verhoeven & van Leeuwe, 2008).

VOCABULARY ASSESSMENT WITH THE PEABODY PICTURE VOCABULARY TEST—REVISED (PPVT-R)

One of the most widely used measures of spoken vocabulary knowledge is the PPVT-R. The test was introduced by Dunn and Dunn (1981) as a measure of receptive vocabulary for children and since then has been widely used in research and practice (e.g., Mahon & Crutchley, 2006; Pankratz, Morrison, & Plante, 2004). The examinee is asked to indicate on a stimulus plate which of four drawings corresponds to a spoken word (noun, verb, or adjective). Due to the special response requirements of the tasks, it is reasonable to expect that perceptual organization and decision-making ability may account for a certain amount of individual variability in performance, although, at least in children, PPVT-R performance loads primarily on verbal comprehension-related factors (Culbert, Hamer, & Klinge, 1989; D'Amato, Gray, & Dean, 1988). The utility of the test has been appraised for a variety of clinical groups and purposes. Ease of administration and scoring, and that the test does not require a verbal response, render it appealing for assessing language as well as general intellectual ability in children who are mentally retarded, severely language impaired, or display emotional and behavioral problems (Ollendick, Finch, & Ginn, 1974; Pasnak, Wilson-Quayle, & Whitten, 1998). In general, standard PPVT-R scores tend to be higher than those on the third edition of the WISC (WISC-III) and Wechsler Preschool and Primary Scale of Intelligence verbal subtests, especially among "special" ability groups (Vance, Kitson, & Singer 1985; Vance, West, & Kutsick, 1989). It has even been reported that the strength of the relation between PPVT-R test scores and achievement tests is somewhat higher than the relation between other verbal tests, characterized by more elaborate response requirements, and the same achievement tests (Ollendick et al., 1974).

There is growing evidence on the strong predictive validity of the PPVT-R for academic skills. For instance, PPVT-R scores have been used as a measure of lexical (word) knowledge in studies investigating sources of individual variability in reading achievement (for a recent review see Joshi, 2005). Word knowledge may account for discrepancies between decoding and reading comprehension skills (Oakhill, Cain, & Bryant, 2003). In this context, it is necessary to identify valid indices of both the depth and the breadth of lexical knowledge (Ouellette, 2006), which, according to one view, should reflect the existence of lexical entries associated with redundant (phonological, orthographic, and semantic) information (Perfetti, 2007; Perfetti & Hart, 2001). In addition to the logical expectation that strong vocabulary knowledge would facilitate fluency and reading comprehension (Dixon, LeFevre, & Twilley, 1988; Frost, Madsbjerg, Niedersoe, Olofsson, &

Sorensen, 2005), several recent empirical findings appear to support this thesis (Vellutino, Tunner, Jaccard, & Chen, 2007; Yovanoff, Duesbery, Alonzo, & Tindal, 2005). For instance, in a recent study two measures of lexical/semantic knowledge (the WISC-III vocabulary test and the PPVT-R) were found to account (both jointly and independently) for a significant amount of variance in passage comprehension scores even after controlling for word-level reading skills, automatized naming ability, and nonverbal intelligence (Protopapas, Sideridis, Mouzaki, & Simos, 2007). A vocabulary composite (made up of the PPVT-R and the Wechsler Abbreviated Scale of Intelligence vocabulary test) was found to account for a significant amount of reading comprehension variance (a composite including the Pearson Individual Achievement Test—Revised print sentence comprehension and Gray Oral Reading Test passage comprehension subtests) after controlling for age, listening comprehension, and nonword reading in young adults with a wide range of reading ability (Braze, Tabor, Shankweiler, & Mencl, 2007).

The present study was designed to assess the growth of receptive vocabulary using the Greek adaptation of PPVT-R in Greek as L2 in school-aged children born in immigrant families of Albanian descent. All children attended monolingual schools starting in Grade 1 and were followed up between Grades 2 and 6. Specific aims of the study were the following: (a) to assess potential differences in vocabulary growth between L1 (Greek children) and L2 (immigrant children) over the 2-year study period, (b) to determine the extent to which growth rate differences can be accounted for by gender and parental education (as a proxy for SES), and (c) to determine the potential effect of nonverbal cognitive ability on vocabulary acquisition in the two groups of children. Individual growth curves were computed using a hierarchical two-level model allowing the estimation of the effects on vocabulary growth rates of ethnic group, parental education, and baseline nonverbal ability.

METHOD

Participants

The full sample comprised 580 children from elementary school Grades 2, 3, and 4, recruited from 17 Greek elementary schools in Crete, Attica (including the Athens metropolitan area), and the Ionian islands. School selection followed a stratified randomized approach in an effort to include units representative of urban (7 schools), rural (3 schools), and semiurban areas (7 schools). Children were selected randomly from each class, but only those whose parents gave written permission for participation in the research were included in the study. All participating students were fluent speakers of Greek, had never been retained in the same grade, and did not suffer from any mental or emotional impairment prohibiting enrollment in regular education according to school records. Fluency in Greek was determined qualitatively by teacher report and was assessed solely in order to identify students with poor basic interpersonal communicative skills as an exclusionary criterion. It should be pointed out that an adequate basic interpersonal communicative skills level may mask difficulties in cognitive/academic language proficiency (including vocabulary, which was the focus of the present study; Baker, 2011; Cummins,

Table 1. *Sample demographic and psychoeducational characteristics at Wave 1*

	Immigrant	Greek Entire Sample	Greek Matched
N	50	530	198
Gender (M/F)	26/24	254/276	95/103
Grade (2/3/4)	19/14/17	185/175/170	71/64/63
Parental education (H/L)	0/50	74/456	0/50
WISC-III			
Block design SS	7.34 ± 2.63	9.46 ± 3.20	7.48 ± 2.27
Vocabulary SS	7.32 ± 2.03	9.63 ± 2.89	8.85 ± 2.65
PPVT-R (*z* score)	–0.88 ± 1.20	0.08 ± 0.93	0.20 ± 0.99

Note: Greek matched, native Greek students matched on average with the group of immigrant students on parental education and block design score at Time 1; M/F, male/female; WISC-III, Wechsler Intelligence Scale for Children, Third Edition; SS, scaled score; PPVT-R, Peabody Picture Vocabulary Test—Revised. There were 14 mixed couples with respect to parental education in this group, and students were categorized according to maternal education.

2000). The sample included 530 monolingual Greek children and 50[1] bilingual children born to immigrant families of Albanian origin. Table 1 presents the demographic characteristics of the sample. Parental education was coded as a categorical variable with two levels (high/low). Children in the high parental education group had both parents with more than 9 years of formal education whereas children in the low parental education group had both parents with 9 or fewer years of formal education. According to school records, there were no "mixed" couples (with respect to education classification) in the immigrant group. There were 21 "mixed" couples in the larger Greek sample of 530. In those cases, students were categorized according to maternal education. Four bilingual students had attended school (1–4 years) in their home country (one fourth grader had completed kindergarten–third grade in Albania, two third graders had attended first and second grades in Albania, and another third grader had attended second grade in his home country). Based on teacher and student report, the language spoken at the students' home was only Greek (12%), mainly Greek (29%), only Albanian (12%), mainly Albanian (12%), or both languages about equally (35%). However, these estimates were largely qualitative in nature and not factored into the statistical models.

Procedure

Children were tested on five consecutive occasions separated by approximately 6 months. At each measurement wave all children were tested individually in two 40-min sessions by a group of undergraduate and graduate students during a period of 3 weeks. Wave 1 was conducted in March of 2005. Examiners underwent rigorous training and were closely monitored by the study coordinators in order to standardize administration procedures.

The Greek version of PPVT-R (Dunn & Dunn, 1981) served as the primary measure. Psychometric properties of the Greek version of the test are reported elsewhere (Simos, Sideridis, Protopapas, & Mouzaki, 2011). Briefly, extensive modifications were deemed necessary in the structure of the test in order to accommodate linguistic and cultural differences. Two of the original plates were excluded (items # 65 and 69), leaving a total of 173 plates, and the target word (and corresponding pictorial stimulus) was altered in 44 of the original stimulus plates. Changes in the order of presentation for several items were also deemed necessary based on item difficulty (percentage correct responses in pilot data). Finally, a more lenient stopping rule (8 failures in 10 consecutive trials) was adopted to increase the sensitivity of the measurements. In cases of an incorrect response within the first 6 items (50–55), the examiner administered items reversely (starting from item 49), until a baseline of 6 consecutive correct responses was reached. Administration started with item 50, which is the starting point for children aged 6.5 years in the English version of PPVT-R. Modeling of group differences (Greeks and immigrants) at the item level using the Rasch model (1980; Smith & Smith, 2004) failed to reveal systematic item bias in favor of either group (Simos et al., 2011).

The assessment battery included word and pseudoword reading accuracy, pseudoword and sight word efficiency (fluency), reading comprehension, rapid automatized naming, and spelling. The WISC-III vocabulary and block design subtests (Greek standardization; Georgas, Paraskevopoulos, Bezevegis, & Giannitsas, 1997) were also administered to the entire sample, using standard procedures. WISC-III standard scores are reported in Table 1 to facilitate comparison with other studies. The block design standard score was also entered as a covariate in some of the analyses reported below in order to assess and control for systematic individual differences in nonverbal cognitive ability at both the intercept and the slope levels. Block design scores show moderate correlations with PPVT-R scores, suggesting that performance on the latter is affected to some extent by nonverbal skills.

Analyses

Individual growth curve analysis was used to investigate change in vocabulary knowledge over time. The dependent variable used in these analyses was the PPVT-R raw score for each participant at each of the five measurement points. Wave (or time) with five levels served as the Level 1 predictor. Two sets of analyses were conducted: unconditional, including only the Level 1 predictor, and conditional. The purpose of the unconditional models was to verify that ample amounts of variance were present around the mean and slope within as well as between students. The general form of the equations describing the unconditional model are the following:

Level 1

$$\text{PPVT-R} = \pi_{0i} + \pi_{1i} \times \text{TIME}_{ij} + e_{ij} + \pi_{2i} \times \text{TIME}_{ij}^2 + e_{ij}. \quad (1)$$

Level 2

$$\pi_{0i} = \beta_{00} + r_0, \tag{2}$$

$$\pi_{1i} = \beta_{10} + r_1, \tag{3}$$

$$\pi_{2i} = \beta_{20} + r_2, \tag{4}$$

with the terms π_0, π_1, and π_2 involving the grand mean of PPVT-R, its linear slope, and quadratic slope, respectively. TIME_{ij} and TIME^2_{ij} represent the linear and quadratic terms, respectively. At Level 2, π_0, π_1, and π_2 are modeled as dependent variables and the terms β_{00}, π_{10}, and π_{20} reflect the intercepts of those estimates.

At Level 1 the equations involved in the conditional and unconditional models were identical. Each conditional model included an additional term serving as a between-subjects Level 2 predictor (ethnic/language group, parental education group, gender, or block design score), according to the following set of conceptual equations:

Level 1

$$\text{PPVT-R} = \pi_{0i} + \pi_{1i} \times \text{TIME}_{ij} + e_{ij} + \pi_{2i} \times \text{TIME}^2_{ij} + e_{ij}.$$

Level 2

$$\pi_{0i} = \beta_{00} + \beta_{01}(\text{PREDICTOR}) + r_0, \tag{5}$$

$$\pi_{1i} = \beta_{10} + \beta_{11}(\text{PREDICTOR}) + r_1, \tag{6}$$

$$\pi_{2i} = \beta_{20} + \beta_{21}(\text{PREDICTOR}) + r_2, \tag{7}$$

with the terms π_0, π_1, and π_2 reflecting the grand mean of PPVT-R, its linear slope (growth), and its quadratic effect, respectively. TIME_{ij} and TIME^2_{ij} represent the linear and quadratic growth terms, respectively. At Level 2 π_0, π_1, and π_2 are modeled as dependent variables and are predicted by each one of the four Level 2 predictors, entered into the model separately. The π_{01}–π_{21} terms reflect partial regression coefficients as in multiple regression. The π_{00}, π_{10}, and π_{20} terms reflect the intercepts of the respective equations. All models involved random effects estimation and modeled each individual's growth trajectory over time.

Level 1 and Level 2 models were fitted to the PPVT-R raw data with restricted maximum likelihood estimation, using Statistical Package for the Social Sciences 16.0 mixed-models analysis. Best fitting models of growth were assessed through a multistage procedure. Unconditional models were tested initially by examining fixed effects for the intercept, (linear) slope, and quadratic parameters. If the fixed slope and/or quadratic effect were significant, corresponding predictors were included in the model to explain the variance in those parameters. Examination of residual plots and comparative fit indices (−2 restricted log likelihood, Akaike information criterion, and Schwartz Bayesian criterion) were used to examine the overall fit of the model (Francis, Schatschneider, & Carlson, 1999). Ethnic/language group, parental education group, and gender were dummy coded with Greek, high education, and male children as the respective reference categories (= 0), whereas Time 1 was set as the reference value for time. Finally, block design scores measured at Time 1 were centered around the grand mean.

RESULTS

Table 2 presents PPVT-R raw and z scores at each measurement point for each group and for the entire sample. Mean block design standard scores were 9.28 ($SD = 3.2$) for the entire sample (9.43, $SD = 3.2$ for native Greek participants and 7.70, $SD = 2.89$ for immigrants). Table 3 presents correlation coefficients between independent variables which did not exceed $r = .2$, suggesting that problems with multicollinearity were not likely. Additional tests, regressing each of the independent variables on the combination of the remaining independent variables, corroborated this conclusion, revealing tolerance values > 0.965 in all cases.

Unconditional analyses

An unconditional growth model was fitted on data from the entire sample under the assumptions of (a) compound symmetry and (b) autoregressive covariance structure. Based on the three criteria of model fit mentioned previously and the estimate of the covariance of residuals across waves was significantly different from 0 ($r = .12$, $z = 2.32$, $p = .02$), the model including autoregressive error variance components provided the best fit to the data and was adopted in further analyses. Each model was then tested with and without assessing intercept and slope random effects as described in Equations 1–3.

As shown in Table 4, random intercept and slope effects (variability of intercepts and slopes between participants, as well as covariation of slopes and intercepts) were significant, justifying further exploration of the sources of between-participant variability in model parameters by adding additional Level 2 predictors. Moreover, inclusion of a quadratic term (according to Equation 4) further improved model fit (based on the difference in −2 log likelihood fit indices, with 4 degrees of freedom, between models differing only in the presence of the quadratic term $\Delta\chi^2 = 101$, $p = .0001$). Linear and quadratic slopes for each group are displayed in Figure 1. Table 3 indicates that all three parameters (intercept, linear, and

Table 2. *PPVT-R raw and age-adjusted z scores by linguistic/ethnic group at five measurement points*

		Time 1	Time 2	Time 3	Time 4	Time 5
PPVT-R raw scores	Greek	116 (17.3)	123 (15.4)	128 (13.9)	132 (13.2)	134 (12.7)
	Immigrant	99 (23.2)	108 (20.5)	115 (19.8)	122 (16.6)	125 (12.7)
	Total	114.5 (18.5)	121.8 (16.4)	126.7 (14.9)	130.9 (13.7)	133.7 (12.9)
PPVT-R z scores	Greek	0.086 (0.93)	0.087 (0.93)	0.068 (0.94)	0.051 (0.97)	0.059 (0.98)
	Immigrant	−0.92 (1.19)	−0.87 (1.18)	−0.78 (1.24)	−0.59 (1.15)	−0.69 (0.94)
	Total	0.00 (0.99)				

Note: PPVT-R, Peabody Picture Vocabulary Test—Revised.

Table 3. *Correlation matrix of predictor and dependent variables*

	1	2	3	4	5	6	7	8	9	10
1. Gender[a]	1									
2. Parental education[b]	.088	1								
3. Ethnic/linguistic group[c]	−.023	.082	1							
4. PPVT-R Time 1	−.039	−.120†	−.256†	1						
5. PPVT-R Time 2	−.017	−.142†	−.260†	.799†	1					
6. PPVT-R Time 3	−.041	−.156†	−.230†	.796†	.799†	1				
7. PPVT-R Time 4	−.035	−.117†	−.177†	.770†	.789†	.802†	1			
8. PPVT-R Time 5	−.067	−.095	−.195†	.730†	.729†	.800†	.793†	1		
9. WISC-III block design	−.075	−.121†	−.136†	.450†	.432†	.443†	.420†	.412†	1	
10. WISC-III block vocabulary	−.026	−.063	−.183†	.664†	.646†	.649†	.644†	.612†	.423†	1

Note: PPVT-R, Peabody Picture Vocabulary Test—Revised; WISC-III, Wechsler Intelligence Scale for Children, Third Edition.
[a]Male.
[b]High Education group.
[c]Native Greek.
†$p = .001$ (statistical significance at the .01 level requires $r > .096$).

Table 4. *Unconditional model for Peabody Picture Vocabulary Test—Revised for the entire sample*

	Fixed Effects				Variance Estimates			
	Estimate	*SE*	*t*	*p*	Estimate	*SE*	Wald *z*	*p*
Intercept	114.56	0.76	150.75	<.0001	300.83	19.95	15.07	<.0001
Slope	7.63	0.37	20.43	<.0001	26.13	6.02	4.34	<.0001
Quadratic	−0.72	0.08	−8.69	<.0001	0.80	0.31	2.56	.01
Residual					38.39			

Note: The intercept represents the initial level of performance (at Time 1), the slope represents the semiannual rate of linear growth, and the quadratic parameter indicates the average degree of reduction in growth rate across the five measurement occasions.

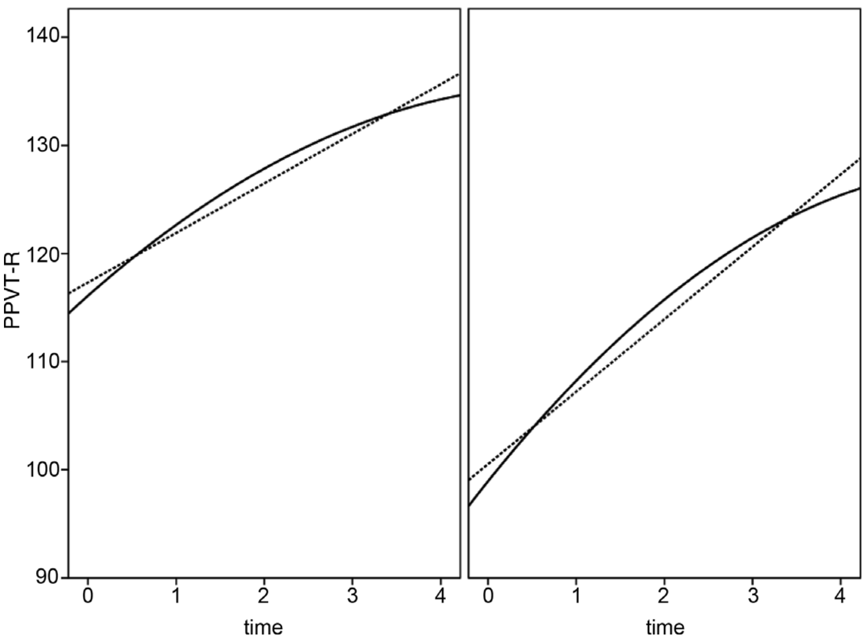

Figure 1. Linear (dotted lines) and quadratic (solid lines) average growth curves for Peabody Picture Vocabulary Test—Revised (PPVT-R) raw scores over five measurement points for (left) native Greek and (right) immigrant students.

quadratic) were valuable estimates of growth in PPVT-R scores. In addition to the expected finding that PPVT-R scores at Time 1 were significantly different from 0, results demonstrated significant linear growth between Time 1 and Time 5, averaging 7.63 points every 6 months, as well as a small but significant reduction in growth rate over time (quadratic effect), which averaged 0.72 points between adjacent measurement points.

Conditional analyses: Separate predictor models

In a second set of analyses we examined potential systematic effects of linguistic/ethnic background, gender, parental education, and initial score on the block design WISC-III subtest on growth parameters. Each model is described by Equations 1 (at Level 1) and 5–7 (at Level 2). For instance, Equation 8 describes the full (mixed) model for ethnic group, including only the linear term:

$$\begin{aligned}\text{PPVT-R} = {} & \beta_{00} + \beta_{01} \times \text{Ethnic Group} + r_0 + \beta_{10} \times \text{TIME}_{ij} \\ & + \beta_{11} \times \text{Ethnic Group} \times \text{TIME}_{ij} + r_1 \times \text{TIME}_{ij} + e_{ij}, \qquad (8)\end{aligned}$$

whereas Equation 9 describes the mixed model for ethnic group by adding the quadratic term:

$$\begin{aligned}\text{PPVT-R} = {} & \beta_{00} + \beta_{01} \times \text{Ethnic Group} + r_0 + \beta_{10} \times \text{TIME}_{ij} \\ & + \beta_{11} \times \text{Ethnic Group} \times \text{TIME}_{ij} + r_1 \times \text{TIME}_{ij} + \beta_{20} \times \text{TIME}^2_{ij} \\ & + \beta_{21} \times \text{Ethnic Group} \times \text{TIME}^2_{ij} + r_2 \times \text{TIME}^2_{ij} + e_{ij}. \qquad (9)\end{aligned}$$

As shown in Table 5, gender did not contribute to the prediction of any of the growth model parameters and was therefore not considered further. In contrast, linguistic/ethnic group entered as a Level 2 predictor improved model fit as indicated by a significant likelihood ratio test with 3 degrees of freedom, $\Delta\chi^2 = 48$, $p = .0001$ (comparing the model described by Equation 1 [Model 1 in Table 5] with the model described by Equation 9 [Model 3 in Table 5]). The fixed effect of linguistic/ethnic group on intercepts was significant ($\hat{\beta} = -17.34$, $SE = 2.61$, $t = -6.64$, $p = .0001$), although the fixed effects of group on the slope ($\hat{\beta} = 2.45$, $SE = 1.35$, $t = 1.81$, $p = .07$) and quadratic trends ($\hat{\beta} = -0.17$, $SE = 0.30$, $t = -0.57$, $p = .57$) were not. Inspection of the random effects table indicated that ethnic group accounted for as much as 8% of the total shared between-participant variance in intercepts, but only for 1% of the total shared between-participant variance in slopes and <0.1% of the total shared between-participant variance in the quadratic component.

Given, however, that the focus of the present study was on individual differences in growth rates of PPVT-R scores, we considered the possibility that models including a quadratic parameter may have obscured group differences in growth rate, especially in view of the relatively small size of the immigrant group. The conditional model with linguistic/ethnic group as a predictor, which only estimated linear growth (see Equation 8, Model 3, in Table 6), provided a better fit compared to the unconditional linear-only growth model (Equation 1, Model 1, in Table 6), as indicated by a significant likelihood ratio test with 2 degrees of freedom ($\Delta\chi^2 = 48$,

Table 5. *Fixed effects and variance estimates for unconditional (Model 1) and conditional growth models (Models 2–5), considering both linear and quadratic growth in Peabody Picture Vocabulary Test—Revised scores*

	Model 1	Model 2	Model 3	Model 4	Model 5
Fixed Effects					
Intercept	114.6****	116.7****	116.1****	122.9****	114.5****
Linear slope	7.63****	7.23****	7.41****	7.55****	7.61****
Quadratic growth	−0.72****	−0.59*	−0.71****	−0.83***	−0.72****
Gender					
Intercept		−1.41			
Linear growth		0.24			
Quadratic growth		−0.09			
Linguistic group					
Intercept			−17.34****		
Linear growth			2.45		
Quadratic growth			−0.17		
Education					
Intercept				−10.48****	
Linear growth				0.03	
Quadratic growth				0.15	
Block design					
Intercept					0.69****
Linear growth					−0.11***
Quadratic growth					0.01
Variance Estimates					
Intercept	300.8****	300.9****	277.8****	284.2****	232.4****
Linear growth	26.1***	26.3****	25.9****	23.9****	24.9****
Quadratic growth	0.80**	0.81**	0.81**	0.74*	0.80**

Note: Significance probabilities of the estimates associated with the corresponding Wald z scores are $*p < .05$, $**p < .01$, $***p < .001$, and $****p < .0001$.

Table 6. *Fixed effects and variance estimates for unconditional (Model 1) and conditional growth models (Models 2–4) that considered only linear growth in Peabody Picture Vocabulary Test—Revised scores*

	Model 1	Model 2	Model 3	Model 4	Model 5
			Fixed Effects		
Intercept	115.8****	117.7****	117.2****	124.5****	115.7****
Linear slope	4.78****	4.94****	4.63****	4.28****	4.78****
Gender					
Intercept		−1.27			
Linear growth		−0.11			
Linguistic group					
Intercept			−17.04****		
Linear growth			1.75****		
Parental education					
Intercept				−10.74****	
Linear growth				0.62	
Block design					
Intercept					0.67****
Linear growth					−0.06****
			Variance Estimates		
Intercept	275.4****	275.5****	253.3****	262.8****	209.9****
Linear growth	3.33****	3.34****	3.12****	3.59****	2.80****

Note: Significance probabilities of the estimates associated with the corresponding Wald z scores are $^{****}p < .0001$.

$p = .0001$). This finding suggested that group was a significant predictor not only of the initial PPVT-R score but also of the semiannual rate of growth of PPVT-R scores. Table 6 further reveals that parental education was a significant predictor of initial PPVT-R scores but not of PPVT-R growth rates. Equation (8) describes this model by substituting parental education for ethnic group. Finally, block design initial score was a significant predictor at both the intercept and the (linear) slope levels, whereas gender failed to contribute to the estimation of any of the growth parameters. Examination of the corresponding random effects tables indicated that ethnic group, parental education, and block design variables accounted for 8%, 5%, and 24%, respectively, of the total shared between-participant variance in intercepts. Ethnic group and block design also accounted for 6% and 16% of the total shared between-participant variance in slopes. Therefore, when entered separately as Level 3 predictors, ethnic group and baseline nonverbal cognitive ability emerged as significant predictors of vocabulary growth. With the exception of gender, which was not associated with significant effects at either the intercept or the slope levels, the joined effect of the remaining three independent variables was examined next.

Conditional analyses: Combined predictor models

The upper section of Table 7 presents results of the combined conditional model, which included time-invariant variables (linguistic/ethnic group, parental education group, and baseline block design scores) as Level 2 predictors.

Linguistic/ethnic group remained a significant predictor of initial vocabulary scores. Controlling for initial block design scores and parental education, immigrant students scored on average 9.84 points lower than native Greek speakers at Time 1. Linguistic/ethnic group also remained a significant predictor of vocabulary growth. Thus, PPVT-R scores increased by an average of 4.65 points every 6 months among Greek students, and immigrant students demonstrated faster growth by 1.1 points (again controlling for the significant effect of initial block design scores on growth rates).

Table 7 further reveals that baseline nonverbal cognitive ability was a significant predictor of both initial PPVT-R score and PPVT-R growth rate independent of the effects of Linguistic/ethnic group and parental education. Finally, the effect of parental education on initial vocabulary estimates persisted even after controlling for baseline block design scores (the effect of parental education on PPVT-R growth rate remained nonsignificant as in the conditional single-variable model of Table 6 (Model 4).

Mixed-models analysis results are presented for each ethnic group separately in the lower sections of Table 7. Here it becomes apparent that parental education did not account for a significant amount of variance in vocabulary growth among native Greek students (as in the case of the entire sample). The effect of parental education was not assessed in the immigrant group because all students had parents in the low education group). Moreover, baseline block design scores were significant predictors of vocabulary initial scores and growth rates for both native Greek students and immigrants.

Table 7. *Predictors of growth in Peabody Picture Vocabulary Test—Revised scores for the entire sample and for each linguistic/ethnic group separately*

	Fixed Effects				Variance Estimates			
	Estimate	*SE*	*t*	*p*	Estimate	*SE*	Wald *z*	*p*
Entire Sample ($N = 580$)								
Intercept	123.87	1.96	62.94	<.0001	195.54	15.5	12.6	<.0001
Ethnic group	−9.84	2.57	−3.83	<.0001				
Parental education	−8.83	2.11	−4.17	<.0001				
Block design	0.63	0.05	10.33	<.0001				
Slope	4.33	0.38	11.44	<.0001	3.01	.72	4.15	<.0001
Ethnic group	1.10	0.51	2.17	.03				
Parental education	0.42	0.41	1.04	.23				
Block design	−0.06	0.01	−4.98	<.0001				
Residual					49.50			
Native Greek Speakers ($N = 530$)								
Intercept	124.07	1.90	65.40	<.0001	176.7	14.8	11.9	<.0001
Parental education	−9.05	2.04	−4.43	<.0001				
Block design	0.61	0.06	10.03	<.0001				
Slope	4.65	0.13	36.27	<.0001	2.89	0.72	3.99	<.0001
Parental education	0.49	0.40	1.21	.23				
Block design	−0.05	0.01	−4.4	<.0001				
Residual					49.37			
Immigrants[a] ($N = 50$)								
Intercept	108.12	21.2	5.10	<.0001	398.6	107.0	3.71	<.0001
Block design	1.00	0.34	2.98	.005				
Slope	7.50	3.20	2.24	.025	3.17	3.98	0.79	.43
Block design	−0.15	0.05	−2.83	.008				
Residual					61.11			

[a]Parental education was not entered into the model because all parents had ≤9 years of formal education.

As a complementary test of the potential mediating effect of the predictor variables on growth parameters, a second set of analyses were performed with linguistic/ethnic group as the sole predictor variable. In this case, however, the reference group included native Greek students matched, on average, with the group of immigrant students on parental education and block design score at Time 1. Demographic and psychoeducational data on the matched subgroup of native Greek students are shown in Table 1. The two groups did not differ on gender and grade distribution ($p > .6$ in both cases) or in block design score ($p > .74$). Results of the individual growth curve analyses indicated that group remained a significant predictor of PPVT-R scores at both the intercept and the slope levels (see Table 8) even after controlling for the initial differences between groups on parental education and block design.

DISCUSSION

The present study reports 2-year longitudinal data on the growth of receptive vocabulary in Greek as L2 in children born in immigrant families from Albania attending Greek schools. Five hundred eighty children aged 6 to 9 years took part in this study, including 530 monolingual Greek speakers and 50 Albanian children with Greek as an additional language. Vocabulary size was measured on the Greek-adapted version of PPVT-R, a test that does not require verbal responses and does not show systematic ethnic or gender bias at the item level, as indicated by Rasch model analyses (Simos et al., 2011). Therefore, differences between Greek and immigrant children on this test may not be attributed to cultural differences in the perception and processing of specific stimuli (such as differences associated with differential exposure to the depicted objects/scenes and concepts) or to differential degree of knowledge for everyday versus academically relevant lexical entities, as discussed by Bialystok and Craik (2010). Results indicated that although immigrant children scored significantly lower than native Greek students at baseline, they showed significantly steeper vocabulary acquisition slopes. Two potential correlates of vocabulary growth were examined systematically as moderators of the impact of ethnic group: parental education and nonverbal cognitive ability (gender was not a significant predictor of vocabulary intercepts or slopes and was not examined further). Although the impact of parental education on baseline vocabulary scores or vocabulary growth could not be assessed in the immigrant group due to lack of adequate variance on this parameter, two of our findings point away from this factor as an independent contributor to vocabulary development among Albanian immigrant children. First, parental education was not related to the rate of growth of PPVT-R scores over the 2-year study period in the entire cohort of students ($N = 580$; although as expected students whose parents had achieved a higher educational level scored significantly higher on PPVT-R at baseline). Second, immigrant students showed significantly faster vocabulary growth rates compared to a group of native Greek students matched on age and parental education.

The second factor that was examined in the present study was nonverbal cognitive ability as measured by scores on the block design WISC-III subtest. This task involves visuospatial processing skills as well as problem solving and construction

Table 8. *Predictors of growth in Peabody Picture Vocabulary Test—Revised scores for the entire matched cohort and for the matched groups of Greek and immigrant students separately*

	Fixed Effects				Variance Estimates			
	Estimate	*SE*	*t*	*p*	Estimate	*SE*	Wald *z*	*p*
Matched Cohort ($N = 248$)								
Intercept	113.22	1.29	87.87	<.0001	296.9	30.79	9.64	<.0001
Ethnic group	−11.70	3.21	−3.65	<.0001				
Slope	4.92	0.23	21.83	<.0001	4.66	1.17	3.96	<.0001
Ethnic group	1.21	0.57	2.13	.034				
Residual					47.37			
Native Greek ($N = 198$)								
Intercept	113.22	1.21	93.76	<.0001	258.79	29.47	8.78	<.0001
Slope	4.92	0.22	22.35	<.0001	4.49	1.19	3.76	<.0001
Residual					45.03			
Immigrants ($N = 50$)								
Intercept	101.53	3.77	26.96	<.0001	494.97	128.62	3.85	<.0001
Slope	6.10	0.58	10.50	<.0001	5.17	4.47	1.16	.25
Residual					62.13			

ability and emerged as a significant predictor of both baseline PPVT-R scores and vocabulary growth rates in the entire sample of students as well as for each linguistic/ethnic group separately. At least at the intercept level (baseline scores), this finding is not surprising given the moderate zero-order correlations typically found between vocabulary and WISC performance scales in student samples (e.g., Craig & Olson, 1991). However, perhaps the most notable present result is that ethnic group differences in both baseline scores and growth rates persisted after controlling for individual differences in block design scores. This conclusion was established both statistically (by entering block design, linguistic/ethnic group, and parental education together as covariates in the same model, see Table 7) and more directly (by comparing immigrant students to a matched subgroup of native Greek students; see Table 8).

Although most immigrant children ($N = 47$) had attended Greek schools since Grade 1, they had significantly lower average baseline PPVT-R scores than native speakers, a difference which was equivalent to approximately 1 *SD*. This observed difference is in concordance with findings from Spanish bilingual students learning English, who demonstrated a gap in vocabulary knowledge (both breadth and depth) that did not diminish between the beginning and end of the school year (August et al., 2005). The average group difference at baseline was only slightly higher than the estimates reported for a large, ethnically mixed group of immigrant children on the newer version of PPVT (PPVT-III; Bialystok, Luk, Peets, & Yang, 2010). In that cross-sectional study, group differences did not vary with age (at least between 8 and 10 years), whereas in our sample the group difference in mean PPVT-R scores declined to approximately 0.64 *SD* points at the end of the study period (2 years). The study period in the present study was lengthier than in the August et al. (2005) study and covered 2 additional years of age than the Bialystok et al. (2010) study. Educational and linguistic implications of these results are discussed below.

The type of growth in PPVT-R scores is crucial for predicting future growth and, consequently, longer-term outcomes. Even at the age range studied here, there were indications of a small, but significant, deceleration in growth rates over time, suggesting that the rate at which the gap between the two groups is closing may slowly decline in subsequent grades. Even if we assume a continuing linear growth at the same rates as those observed during the current study period, immigrant children would attain similar average PPVT-R scores in no less than 3 to 5 years (depending on the characteristics of the group of Greek children they are compared to). These findings are in complete agreement with the findings from studies with Spanish/English bilingual students. According to a recent synthesis of research evidence, it seems that oral proficiency develops over time, and it takes 3 to 5 years for English-language learners to achieve advanced levels (Genesee et al., 2005). These results have implications for assessment strategies for bilingual, immigrant students as well as for predictions regarding their academic achievement. Many experts argue that one should avoid interpreting low scores on language and language-related achievement tests by immigrant children as indicative of deficits in underlying skills (Cobo-Lewis, Pearson, Eilers, & Umbel, 2002). While not representative of their “true abilities,” results on tests in L2 may nevertheless predict future academic achievement in the host country educational system.

Performance on these tests provides an estimate of the distance between the student's proficiency in a particular aspect of L2 and that expected for his/her monolingual peers. The present data further imply that this distance is covered over time but at a very slow pace, even for children who attended Greek schools since Grade 1. Given the well-documented link between vocabulary and reading comprehension (Oakhill et al., 2003; Yovanoff et al., 2005), these findings suggest that this group of young bilingual children is at risk for academic underachievement.

Reduced performance on vocabulary tests by immigrant children should not be interpreted solely as the result of bilingualism but should also take into account the wider social and economic context in which these students develop. There are numerous reports on the crucial role of SES and parental education on vocabulary acquisition and literacy development in monolingual children (Bowey, 1995; Fenson et al., 1994; Hart & Risley, 1995; Heath, 1983; Morisset, Barnard, Greenberg, Booth, & Spieker, 1990; White, 1982). The relative degree of immersion in L1 and L2 (in the family and at the school) appears to interact in complex ways with SES to determine language proficiency (Cobo-Lewis et al., 2002). Although an effort was made in the present study to control for socioeconomic influences and general cognitive abilities of the children, as potential correlates of initial vocabulary attainment and rate of acquisition, SES and IQ indices may only be used as proxy indices for language capacity in L1. Consequently, little can be said regarding the degree and nature of potential interactions between L1 and L2 at the lexical and semantic levels.

Further, the present study relates to a specific type of bilingualism, resulting from immigration, and may not be generalizable to all forms of bilingualism. In this group of children, bilingualism is typically confounded by other factors (SES and parental education). All immigrant children probably came from low SES families, as indicated by low parental education (fewer than 9 years of formal schooling in Albania). Nevertheless, our finding concerning differential growth rates among native Greek and immigrant students appeared to be largely independent of concurrent group differences in parental education per se, a conclusion supported by two lines of evidence: the lack of a parental education effect on vocabulary growth rates for either group and the native–immigrant difference persisted even after matching the two groups on parental education. It should be noted, however, that parental education, as quantified by years of formal schooling, may not be the best way to match the two groups, because it does not take into account differences between groups in the degree of association between parental education and Greek language ability. The same conclusion can be drawn on the potentially confounding effect of general cognitive ability, which in the present study was measured by a single task (block design subtest).

Unfortunately, detailed information on the relative amount and quality of linguistic experience in Albanian and Greek was not available for all the children in our study, neither were formal measures of L1 vocabulary, mainly because there is no Albanian vocabulary assessment test. Therefore, no conclusions can be drawn about bilingualism per se. In this respect, classifying Greek and Albanian as L1 and L2, respectively, is only tentative in the present study. Some general ethnolinguistic considerations may be useful in this context, however. About 60% of the total population of migrants in Greece come from neighboring Albania (Kasimis & Kasimis, 2004). Albanian immigration started in the early 1990s, and thus a second generation of Albanian immigrants currently lives in Greece. Moreover, Albanian immigrant students in Greek schools experience social exclusion and are the target of a number of negative stereotypes, which may serve as a powerful motive for cultural and linguistic assimilation (Gogonas, 2009). Negative ethnic stereotypes very likely generalize toward the immigrants' L1 in the form of negative attributed value, creating a form of attitudinal "competition" between L1 and L2 (Lambert's [1977] "subtractive" bilingualism). Furthermore, issues of intercultural education and L1 development have little influence in the mainstream educational system (Gkaintartzi & Tsokalidou, 2011). Minority children are mixed together with children having Greek as their native language and attend unsystematic "submersion" classes where L1 does not appear in the school curriculum. This situation may operate as an additional, yet powerful motive to acquire L2 proficiency, at least at the conversational level. Quantitative and qualitative data drawn from second-generation adolescents of Albanian origin on patterns of language use within Albanian households indicate that the Albanian ethnolinguistic group is undergoing rapid language shift (Chatzidaki, 2005; Gogonas, 2009). This notion is in agreement with qualitative data suggesting that Greek had been adopted, at least partially, in everyday communication by a significant portion of immigrant families participating in the present study. In this respect, the Albanian immigrant population is somewhat atypical in comparison to other ethnolinguistic groups of immigrants living in Greece with more pronounced cultural identity features (Polish, Bulgarian, Pakistani, Chinese, etc.). In addition, language contact (Albanian Greek) is spread at least in some parts of the population on each side of the border given that there is an indigenous Greek minority in Albania, and therefore Greek language is spoken within the Albanian state, and a variety of Albanian (Arvanitika) is spoken by an indigenous population in Greece.

These issues are also relevant to the type of bilingualism immigrant children who took part in the present study were likely to have experienced. The majority of immigrant children were very likely exposed to (some) Greek from early on, although the precise timing of their introduction to the Greek language cannot be precisely determined. However, because there were no ethnically mixed couples among the families of immigrant children, we may assume that they were introduced to Albanian before Greek. For some of these children, systematic exposure to conversational Greek probably took place in daycare or preschool, where Greek would be the main language of teacher–student interaction. Therefore, it appears more likely that the language-acquisition phenomenon approached in the present study is consecutive/sequential bilingualism rather than bilingual L1 acquisition (simultaneous bilingualism; De Houwer, 2005; Meisel, 1990).

Motivation for sociocultural (including linguistic) assimilation and early contact with the Greek language and culture by Albanian immigrant children may also account at least in part for the observed accelerated vocabulary growth. Another factor that deserves further study concerns the moderating role of the morphological structure of the Greek language in enhancing metalinguistic awareness in general. Most of the work done in the field of morphology and its potential impact on vocabulary acquisition concerns languages such as English, in which inflectional morphology plays a rather marginal role (e.g., Marchman, Martínez-Sussmann, & Dale, 2004). Greek is a highly synthetic language: The ending of each word is an amalgam morpheme (it contains information about grammatical categories, such as person, number, case, tense, and voice). There is no clear-cut distinction between regular and irregular inflection, but rather, several inflectional patterns varying in regularity. Therefore, it is hard for a nonnative speaker to disentangle the components of the bipartite word structure: stem + inflectional ending. Formal grammar teaching, which in Greek schools is focused mainly on morphology, is likely to promote the development of metalinguistic awareness (for a similar proposal for French, see Duncan, Casalis, & Cole, 2009). This process may further interact with the heightened postulated general metacognitive ability among bilingual children (e.g., Bialystok & Majumder, 1998) to boost vocabulary growth.

To conclude, the findings of this study can be potentially useful to practitioners and can serve as a basis for further investigation, but they must be treated cautiously not only because of the specific characteristics of the sample but also because only a single measure of language proficiency was examined (receptive vocabulary), which naturally fails to capture the entire range of communicative and academic language skills on which immigrant and Greek children may differ.

NOTE

1. As a thoughtful reviewer suggested, the unequal sample sizes could potentially be problematic. We addressed the two possible problems that could arise from this inequality. One reflects the validity of the estimated parameters for the immigrant group and the second the possibility that variability between groups would be greatly different (heteroscedasticity). For the first issue, we bootstrapped the mean estimates of the immigrant and Greek groups (see Efron, 1979, 1982, 1985) using 1,000 replications and identified a bias between the sample's point estimate and that of the bootstrap distribution equal to 0.01 for the Greek group and even less (0.001) for the immigrant group. Thus, the bias in means was negligible and was also reflected by very tight confidence intervals (within ±5 raw points on the PPVT-R). With regard to heteroscedasticity, we compared the variances of the two groups on the dependent variable over time. Results, using the Mauchly test of sphericity, indicated that at Times 1 and 2, variances were equal; at Times 3, 4, and 5, they were unequal and we corrected the degrees of freedom for that effect. This latter effect, however, was indicative of excessive power levels (Cohen, 1992).

REFERENCES

August, D., Carlo, M., Dressier, C., & Snow, C. E. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research and Practice*, *20*, 50–57.

August, D., Carlo, M., Lively, T., Lippman, D., McLaughlin, B., & Snow, C. (1999). *Enhancing vocabulary knowledge and reading comprehension in English language learners.* Paper presented at the Annual Meeting of the American Educational Research Association, Montreal.

August, D., & Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Mahwah, NJ: Erlbaum.

August, D., Snow, C., Carlo, M., Proctor, C. P., De San Francisco, A. R., Duursma, E., et al. (2006). Literacy development in elementary school second-language learners. *Topics in Language Disorders*, *26*, 351–364.

Baker, C. (2011). *Foundations of bilingual education and bilingualism*. Clevedon: Multilingual Matters.

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

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition.* New York: Cambridge University Press.

Bialystok, E., & Craik, F. I. (2010). Cognitive and linguistic processing in the bilingual mind. *Current Directions in Psychological Science, 19*, 19–23.

Bialystok, E., Luk, G., & Kwan, E. (2005). Bilingualism, biliteracy, and learning to read: Interactions among languages and writing systems. *Scientific Studies of Reading, 9*, 43–61.

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition, 13*, 525–531.

Bialystok, E., & Majumder, S. (1998). The relationship between bilingualism and the development of cognitive processes in problem solving. *Applied Psycholinguistics, 19*, 69–85.

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

Bialystok, E., & Shapero, D. (2005). Ambiguous benefits: The effect of bilingualism on reversing ambiguous figures. *Developmental Science, 8*, 595–604.

Bowey, J. A. (1995). Socioeconomic status differences in preschool phonological sensitivity and first-grade reading achievement. *Journal of Educational Psychology, 87*, 476–487.

Braze, D., Tabor, W., Shankweiler, D., & Mencl, E. (2007). Speaking up for vocabulary: Reading skill differences in young adults. *Journal of Learning Disabilities, 40*, 226–243.

Chatzidaki, A. (2005). Modela diglosikis siberiforas se ikogenies Alvanon mathiton. Dedomena apo embiriki erevna [Patterns of bilingual behavior in families of Albanian pupils: An empirical study]. *Epistimes Agoyis*, 79–102.

Cobo-Lewis, A. B., Pearson, B. 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: Multilingual Matters.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159.

Craig, R. J., & Olson, R. E. (1991). Relationship between Wechsler scales and Peabody Picture Vocabulary Test—Revised scores among disability applicants. *Journal of Clinical Psychology, 47*, 420–429.

Culbert, J. P., Hamer, R., & Klinge, V. (1989). Factor structure of the Wechsler Intelligence Scale for Children—Revised, Peabody Picture Vocabulary Test, and Peabody Individual Achievement Test in a psychiatric sample. *Psychology in the Schools, 26*, 331–336.

Cummins, J. (1976). The influence of bilingualism on cognitive growth: A synthesis of research findings and explanatory hypotheses. *Working Papers on Bilingualism, 9*, 1–43.

Cummins, J. (1978). Bilingualism and the development of metalinguistic awareness. *Journal of Cross-Cultural Psychology, 9*, 131–149.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49*, 222–251.

Cummins, J. (1984). *Bilingual education and special education: Issues in assessment and pedagogy*. San Diego, CA: College Hill.

Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire*. Clevedon: Multilingual Matters.

D'Amato, R. C., Gray, J. W., & Dean, R. S. (1988). Construct validity of the PPVT with neuropsychological, intellectual, and achievement measures. *Journal of Clinical Psychology, 44*, 934–939.

De Houwer, A. (2005). Early bilingual acquisition: Focus on morphosyntax and the separate development hypothesis. In J. Kroll & A. De Groot (Eds.), *The handbook of bilingualism* (pp. 30–48). Oxford: Oxford University Press.

Dixon, P., LeFevre, J.-A., & Twilley, L. C. (1988). Word knowledge and working memory as predictors of reading skill. *Journal of Educational Psychology, 80*, 465–472.

Duncan, L., Casalis, S., & Cole, P. (2009). Early metalinguistic awareness of derivational morphology: Observations from a comparison of English and French. *Applied Psycholinguistics, 30*, 405–440.

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

Efron, B. (1979). Bootstrap methods: Another look at the Jackknife. *Annals of Statistics, 7*, 1–26.

Efron, B. (1982). *The jackknife, bootstrap, and other resampling plans*. Philadelphia, PA: SIAM.

Efron, B. (1985). Bootsrap confidence intervals for a class of parametric problems. *Biometrica, 72*, 45–48.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5, Serial No. 242).

Francis, D. J., Schatschneider, C., & Carlson, C. D. (1999). Introduction to individual growth curve analysis. In D. Drotar (Ed.), *Handbook of research methods in pediatric and clinical child psychology* (pp. 51–73). New York: Kluwer Academic/Plenum Press.

Frost, J., Madsbjerg, S., Niedersoe, J., Olofsson, A., & Sorensen, P. M. (2005). Semantic and phonological skills in predicting reading development: From 3–16 years of age. *Dyslexia, 11*, 79–92.

Genesee, F., Lindholm-Leary, K., Saunders, W., & Christian, D. (2005). English language learners in U.S. schools: An overview of research findings.*Journal of Education for Students Placed at Risk, 10*, 363–385.

Georgas, D. D., Paraskevopoulos, I. N., Bezevegis, I. G., & Giannitsas, N. D. (1997). *Ελληνικό WISC–III: Wechsler κλίμακες νοημοσύνης για παιδιά* [Greek WISC–III: Wechsler Intelligence Scales for Children]. Athens: Ellinika Grammata.

Gkaintartzi, A., & Tsokalidou, R. (2011). "She is a very good child but she doesn't speak": The invisibility of children's bilingualism and teacher ideology. *Journal of Pragmatics, 43*, 588–601.

Gogonas, N. (2009). Language shift in second-generation Albanian immigrants in Greece. *Journal of Multilingual and Multicultural Development, 30*, 95–110.

Goodz, N. (1994). Interactions between parents and children in bilingual families. In F. Genesee (Ed.), *Educating second-language children: The whole child, the whole curriculum, the whole community*. Cambridge: Cambridge University Press.

Hakuta, K., Butler, Y. G., & Witt, D. (2000). *How long does it take English learners to attain proficiency.* University of California, Linguistic Minority Research Institute, Policy Report 2000-1. Retrieved from http://www.stanford.edu/hakuta/Docs/HowLong.pdf

Hamers, J. F., & Blanc, M. H. A. (2000). *Bilinguality and bilingualism* (2nd ed.). Cambridge: Cambridge University Press.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H. Brooks.

Heath, S. B. (1983). *Ways with words*. Cambridge: Cambridge University Press.

Jean, M., & Geva, E. (2009). The development of vocabulary in English as a second language children and its role in predicting word recognition ability. *Applied Psycholinguistics, 30*, 153–185.

Joshi, R. M. (2005). Vocabulary: A critical component of comprehension. *Reading and Writing Quarterly, 21*, 209–219.

Kasimis, C., and Kasimis, C. (2004). Greece: A history of migration. *Migration Information Source.* Retrieved from http://www.migrationinformation.org/Profiles/display.cfm?ID=228

Lambert, W. E. (1977). Effects of bilingualism on the individual: Cognitive and sociocultural consequences. In P. A. Hornby (Ed.), *Bilingualism: Psychological, social, and educational implications* (pp. 15–28). New York: Academic Press.

Leseman, P. P. M. (2000). Bilingual vocabulary development of Turkish preschoolers in the Netherlands. *Journal of Multilingual and Multicultural Development, 21*, 93–112.

Macnamara, J. (1966). *Bilingualism and primary education*. Edinburgh: Edinburgh University Press.

Mahon, M., & Crutchley, A. (2006). Performance of typically developing school-age children with English as an additional language on the British Pictures Vocabulary Scales II. *Child Language Teaching and Therapy, 22*, 333–351.

Marchman, V. A., Martínez-Sussmann, C., & Dale, P. S. (2004). The language-specific nature of grammatical development: Evidence from bilingual language learners. *Developmental Science, 7*, 212–224.

Meisel, J. M. (1990). *Two first languages: Early grammatical development in bilingual children*. Dordrecht: Foris.

Morisset, C. M., Barnard, K. E., Greenberg, M. T., Booth, C. L., & Spieker, S. J. (1990). Environmental influences on early language development: The context of social risk. *Development and Psychopathology, 2*, 127–149.

Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes, 18*, 443–468.

Ollendick, T. H., Finch, Jr., A. J, & Ginn, F. W. (1974). Comparison of Peabody, Leiter, WISC, and academic achievement scores among emotionally disturbed children. *Journal of Abnormal Child Psychology, 2*, 47–51.

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98*, 554–566.

Páez, M., & Rinaldi, C. (2006). Predicting English word-reading skills for Spanish-speaking students in first grade. *Topics in Language Disorders, 26*, 338–350.

Pankratz, M., Morrison, A., & Plante, E. (2004). Difference in standard scores of adults on the Peabody Picture Vocabulary Test (revised and 3rd ed.). *Journal of Speech, Language, and Hearing Research, 47*, 714–718.

Pasnak, R., Wilson-Quayle, A., & Whitten, J. (1998). Mild retardation, academic achievement, and Piagetian or psychometric tests of reasoning. *Journal of Developmental and Physical Disabilities, 10*, 23–33.

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*, 357–384.

Perfetti, C. A., & Hart, L. (2001). The lexical bases of comprehension skill. In D. S. Gorien (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity* (pp. 67–86). Washington, DC: American Psychological Association.

Protopapas, A., Sideridis, G. D., Mouzaki, A., & Simos, P. G. (2007). Development of lexical mediation in the relation between reading comprehension and word reading skills in Greek. *Scientific Studies of Reading, 11*, 165–197.

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press.

Rhodes, R. L., Ochoa, S. H., & Ortiz, S. O. (2005). *Assessing culturally and linguistically diverse students: A practical guide*. New York: Guilford Press.

Selinker, L. (1972). Interlanguage. *IRAL, 10*, 209–231.

Simos, P. G., Sideridis, G., Protopapas, A., & Mouzaki, A. (2011). Psychometric characteristics of a receptive vocabulary test for Greek elementary students. *Assessment for Effective Intervention, 37*, 34–49.

Smith, E. V., & Smith, R. M. (2004). *Introduction to Rasch measurement: Theory, models, and applications*. Maple Grove, MN: JAM Press.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Tucker, G. R. (1977). The linguistic perspective. In G. R. Tucker (Ed.), *Bilingual education: Current perspectives: Vol. 2. Linguistics*. Washington, DC: Center for Applied Linguistics.

Uccelli, P., & Páez, M. M. (2007). Narrative and vocabulary development of bilingual children from kindergarten to first grade: Developmental changes and associations among English and Spanish skills. *Language, Speech, and Hearing Services in Schools, 38*, 225–236.

Vance, B., Kitson, D., & Singer, M. G. (1989). Relationship between the standard scores of Peabody Picture Vocabulary Test—Revised and Wide Range Achievement test. *Journal of Clinical Psychology, 41*, 691–693.

Vance, B., West, R., & Kutsick, K. (1985). Prediction of Wechsler preschool and primary scale of intelligence IQ scores for preschool children using the Peabody Picture Vocabulary Test—Revised and the Expressive One-Word Picture Vocabulary Test. *Journal of Clinical Psychology, 45*, 642–644.

Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., & Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. *Scientific Studies of Reading, 11*, 3–32.

Verhoeven, L., & van Leeuwe, J. (2008). Prediction of the development of reading comprehension: A longitudinal study. *Applied Cognitive Psychology, 22*, 407–423.

White, K. R. (1982). The relationship between socioeconomic status and academic achievement. *Psychological Bulletin, 91*, 461–481.

Yovanoff, P., Duesbery, L., Alonzo, J., & Tindal, G. (2005). Grade-level invariance of a theoretical causal structure predicting reading comprehension with vocabulary and oral reading fluency. *Educational Measurement: Issues and Practice, 24*, 4–12.