HHS Public Access
Author manuscript
Biling (Camb Engl). Author manuscript; available in PMC 2015 March 04.

Receptive vocabulary differences in monolingual and bilingual children*

Ellen Bialystok,
York University, Rotman Research Institute of Baycrest

Gigi Luk,
Rotman Research Institute of Baycrest

Kathleen F. Peets, and
York University

Sujin Yang
York University, Tyndale University College

Abstract

Studies often report that bilingual participants possess a smaller vocabulary in the language of testing than monolinguals, especially in research with children. However, each study is based on a small sample so it is difficult to determine whether the vocabulary difference is due to sampling error. We report the results of an analysis of 1,738 children between 3 and 10 years old and demonstrate a consistent difference in receptive vocabulary between the two groups. Two preliminary analyses suggest that this difference does not change with different language pairs and is largely confined to words relevant to a home context rather than a school context.

Research comparing monolingual and bilingual children on a wide variety of linguistic and cognitive tasks often includes a measure of receptive vocabulary, usually the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997). The test is simple to administer, and because it has been standardized on a large population, there is confidence in both the reliability and validity of the results (Bracken & Murray, 1984). The task requires children to point to one of four pictures that best represents a word spoken by the experimenter. The items become increasingly difficult and detailed tables convert children's raw scores to standard scores based on their age. The test has been standardized on an American sample ranging in age from 3 to 89 years old and has a reported population mean of 100 and a standard deviation of 15.

It is not surprising that bilingual children know fewer words in each language than monolingual speakers of one of the language (e.g., Oller, Pearson & Cobo-Lewis, 2007).

Address for correspondence: Ellen Bialystok, Department of Psychology, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada, ellenb@yorku.ca.

Bilingual children need to distribute their language-learning time across two languages, and it is likely that some words occur in a context in which they only use one of their languages. In this sense, there is little reason to think that bilingual children are compromised in their expressive ability and every possibility that their combined vocabulary is equivalent to or greater than the vocabulary of monolingual children.

There is a crucial difference, however, between the vocabulary available for conversational uses of language and the vocabulary that is the basis for the language of schooling. A large body of evidence shows that vocabulary size is a significant predictor of academic achievement and literacy acquisition (Adams, 1990; Kastner, May & Hildman, 2001; Ouellette, 2006; Ricketts, Nation & Bishop, 2007; Rohde & Thompson, 2007; Swanson, Rosston, Gerber & Solari, 2008). In a study of almost 200 ten-year-olds, Smith, Smith and Dobbs (1991) reported significant correlations between PPVT scores and the reading, spelling, and arithmetic subtests of the Wide Range Achievement test (WRAT-R; Jastek & Wilkinson, 1984). Therefore, if bilingual children have a smaller vocabulary than monolingual children in the language of schooling, there may be consequences for the success of those children in school-related assessments. It is difficult to generalize from individual studies to conclusions about relative vocabulary size because each individual study focuses on a small group of children of a particular age. The purpose of the present study is to establish the extent to which the vocabulary differences in the language of schooling (i.e., English) found between monolingual and bilingual children in individual studies reflects a general pattern that applies to a large number of children and a wide range of ages.

Two methods have been used to conduct large-scale analyses that combine data from multiple studies. The first is a meta-analysis in which the individual study that contributed the data is a factor in the analysis. This method is generally used to determine properties of an effect that have been found in a number of studies, typically because the experimental procedures are somewhat different and the overall reliability of the effect needs to be established. The second method is to use an aggregate analysis in which data obtained from different studies are combined and analyzed in a single model without attention to their origin. This method is appropriate for understanding the generalizablity of an effect in a large-scale analysis when the studies contributing the data have used identical procedures. In aggregated datasets, sample heterogeneity increases and if the hypothesized effect persists, the effect is deemed robust and issues of biased sampling are diminished (Curran & Hussong, 2009). In the present case, data obtained on PPVT scores from a large number of studies were combined and submitted to an aggregate analysis. Because the PPVT is a standardized test, the administration was identical in all the studies, so between-study variance is not relevant.

The present study addresses three issues important for understanding the language development and academic achievement of bilingual children in terms of the frequently reported vocabulary deficit in the language of schooling. The first is to determine the extent to which the results of the individual studies are sustained in a large-scale analysis that includes children of a wide range of ages. The second is to investigate whether the non-English language of the bilingual children, classified roughly as East Asian or non-Asian,

interacts with the patterns. The third is to use an item analysis to allow a preliminary investigation for a more precise evaluation of vocabulary differences in the language of schooling. All the children in the present analysis were bilingual in English and another language and all the children were being educated in English. Therefore, differences in English vocabulary are potentially relevant for successful outcomes in academic achievement and literacy.

Method

Participants

PPVT standard scores from a total of 1,738 children who ranged in age from 3 to 10 years old were included in the analysis. All these children had participated in studies conducted by the first author over a period of five years. Of this sample, 772 children were English monolingual speakers and 966 were bilingual speakers. The numbers are not equal because some studies included several groups of bilingual children and only one group of monolingual children (e.g., Bialystok, Luk & Kwan, 2005) and other studies examined only bilingual children (e.g., Luk & Bialystok, 2008).

Although the bilingual sample was recruited from multiple studies, they all satisfied the same selection criteria: they were being educated in English at school, they spoke a non-English language at home with family members, their parents reported that they were fluent in both English and the non-English language, and parents stated that these children used both languages on a daily basis. Children who were learning English as a second language were excluded from the analysis.

Peabody Picture Vocabulary Test (PPVT-III; Dunn & Dunn, 1997)

In the PPVT-III, the experimenter shows the child an easel with four black-and-white line drawings on each page. Children are asked to indicate the picture that matches the word spoken by the experimenter. Younger children are allowed to respond by pointing to the picture, and older children respond by either saying the number corresponding to the picture or by pointing to it. The experimenter records the child's responses to each item.

The test contains 204 trials that are grouped into 17 sets of 12 items each. Items begin with common concrete objects (e.g., a bus) or simple actions (e.g., fly) and become increasingly difficult such that the last sets of items are uncommon objects or abstract concepts (e.g., embossed, vitreous, lugubrious). The starting set is determined by the child's age. A basal set with one or no errors among the 12 items in the set needs to be established. If a child has more than one error in the first set, then the experimenter moves to the first item of the previous set. After establishing a basal set, the experimenter continues testing until the child commits eight or more errors in a set, establishing the ceiling set. The raw score is obtained by subtracting the number of errors from the last item of the ceiling set. This raw score is converted to a standardized score from a table providing age-corrected normative scores.

Results

The standard scores for the monolingual and bilingual children by age are presented in Figure 1. A two-way ANOVA for age group and language group showed significant effects of age, $F(7,1717) = 6.17$, $p < .0001$, and language, $F(1,1717) = 130.31$, $p < .0001$, with no interaction, $F(7,1717) = 1.87$, *ns*. The effect of language is clear, with monolinguals outperforming bilinguals at every age comparison (all *p*s < .001), but the effect of age group is less robust. Using Scheffé contrasts, none of the inter-age group comparisons was significant, but using the less conservative Tukey test for contrasts, there was a significant effect in which 4-year-olds and 5-year-olds (collapsed across language group) obtained higher PPVT scores than 7-year-olds, $p < .05$. Since the effect is small and anomalous, it is possibly a reflection of sampling variance between the children in these two age groups.

The difference in mean score by language group was examined in more detail by plotting the distribution of scores for each group. The lower mean scores for the bilinguals might have been caused by outlier scores that were very low, producing a skewed function, or by a subset of children who obtained low scores for some other reason, producing a bimodal distribution. However, as shown in Figure 2, both distributions produced normal curves, with the overall mean of the monolinguals (M = 106.8, SD = 12.3) being slightly above the population mean ($\mu = 100$) and that of the bilinguals, slightly below (M = 96.3, SD = 13.0). Importantly, both group means are within the normal range of one standard deviation for the population mean. With the distributions approaching normality, it is also possible to examine the additional central tendency measures besides the means. The median and mode for monolinguals were 106.5 and 105 respectively while the same statistics for the bilinguals were 97 and 94. As shown in these statistics, there was a consistent nine-point difference in central tendency measures between monolinguals and bilinguals.

The studies that contributed data to the aggregate analysis reported in Figure 1 included all monolingual and bilingual children, but the bilingual children spoke a variety of non-English languages. With a large sample, it is possible to determine whether children learning English from different language backgrounds, different cultures, and different social histories show different levels of progress. Again, each individual study contains too few children from each group to provide authoritative evidence, but the aggregate analysis allows us to address that issue.

Information about the non-English language was available for approximately two-thirds of the bilingual children. As a preliminary attempt to distinguish between linguistic and cultural backgrounds, children were divided into those who spoke an East Asian (n = 329) or non-Asian (n = 247) language and compared to the monolinguals.[1] There were two reasons for using this non-typological distinction to create the subgroups. The first was that by using a broad criterion, all the bilingual children for whom a non-English language was specified could be included in the analysis, maximizing the sample size. Furthermore, this distinction

---

[1]The East Asian languages included Cantonese, Japanese, Korean, Mandarin, Thai, and Shanghainese. The non-Asian languages included Amharic, Arabic, Croatian, Farsi, French, German, Greek, Gujarati, Hebrew, Hindi, Hungarian, Italian, Kannada, Macedonian, Marathi, Persian, Polish, Portuguese, Punjabi, Romanian, Russian, Serbian, Spanish, Tamil, Telugu, Turkish, Tagalog, and Urdu.

produced subgroups of somewhat comparable numbers. The second was that some researchers have recently suggested that East Asian culture, rather than bilingualism, is responsible for the performance differences reported in studies comparing monolingual and bilingual children (e.g., Lewis, Koyasu, Oh, Ogawa, Short & Huang, 2009). Therefore, it is important to demonstrate that on this linguistic variable at least there is no difference between East Asian bilingual children and other bilinguals.

Since the English monolinguals outnumbered those in the other two groups, 95% confidence intervals (CI) were constructed for each language group to indicate the range of values around the group mean that can be determined with 95% confidence. These values were used as post-hoc comparisons. A CI of 0 indicates no differences between groups. The results are shown in Figure 3. Language group had a significant effect on PPVT scores, $F(2,1348) = 103.13$, $p < .0001$. The 95% confidence intervals of the difference between means indicated that there was no difference between the two bilingual groups (−3.4, 1.2) but a significant difference between each of them and the monolinguals [Non-Asian: (−7.8, −11.5); East Asian: (−12.4, −9.1)].

The third question addresses the difference between vocabulary that supports conversational uses of language and school-based academic achievement. The words in the PPVT cover a wide range of topics, parts of speech, and contexts. Since the bilingual children primarily used English at school and the non-English language at home, it is possible that these contexts selectively disadvantage certain portions of English vocabulary. Therefore, an item analysis was conducted to classify words on the basis of their primary context being home or school. These categories are not absolute as most linguistic repertoire is used in all contexts, but a bias for higher proficiency in a subset of the tested receptive vocabulary might help to understand children's functional use of language, especially in school. Criteria for inclusion in the home category were as follows: food and household items (e.g., *squash*, *camcorder*, *pitcher*), culture-specific items (e.g., *canoe*, *camper*) and words that were unlikely to occur in a classroom context (e.g., *horrified*). Criteria for inclusion in the school category were as follows: professions (e.g., *astronaut*), animals or plants (e.g., *raccoon*), shapes (e.g., *rectangle*), musical instruments (e.g., *harp*), and words reflecting school experiences (e.g., *writing*) that were more associated with school activities and discussion. Using these criteria, two postdoctoral fellows independently classified all the items from sets 1–10 of the test; inter-rater raw agreement was 91.7%, and chance corrected agreement using Cohen's Kappa was .73, which would be interpreted as satisfactory inter-rater reliability. Consensus was reached on all disagreements to arrive at a categorization of each word in the test as either "home" (total of 24 items from sets 1–10) or "school" (total of 96 items from sets 1–10) based.

An analysis of receptive vocabulary in terms of these subset scores was applied to 161 children from the larger sample who were aged 6;0–6;11. The mean percentage correct in each category was tabulated for each participant, producing a score for each of the home word and school word categories. These percentages by category were analyzed to determine whether the context of the vocabulary item interacted with language group.

The results of this analysis by word context are shown in Table 1. There was no age difference between the monolingual and bilingual children in the subsample, but as in the larger analysis, the overall PPVT score was higher for monolinguals than for bilinguals, $F(1,160) = 8.99$, $p < .003$. However, a two-way ANOVA comparing the percentage of correct responses in each category for the two language groups indicated effects of language group, $F(1,160) = 13.42$, $p < .0003$, word type, $F(1,160) = 8.49$, $p < .004$, and their interaction, $F(1,160) = 5.53$, $p < .02$. Monolinguals knew more words than bilinguals in the home category, $F(1,160) = 12.38$, $p < .0006$, but performance of the two groups was more comparable for school words, $F(1,160) = 3.74$, $p = .06$.

Discussion

Using a large sample of children aged between 3 and 10 years, the mean standard score on the PPVT was significantly lower for bilinguals than for monolinguals in each age group. Thus, individual studies that report lower vocabulary scores for bilingual children than for monolinguals are reflecting a general pattern in which bilingual children tend to know fewer words in one of their languages than comparable monolingual speakers of that language. Importantly, the distribution of scores was normal for both language groups, and the difference in mean score reflected a small shift in the central tendency. Thus, it is unlikely that the lower mean score of the bilingual group was driven by a subgroup of bilinguals based on such possible factors as non-English language. Instead, the distribution presented in Figure 2 indicates that many bilinguals achieved higher scores than monolinguals on this test, but that the overall mean signifies a general tendency for bilinguals to have smaller vocabularies in each language. It might have been expected that the gap between monolingual and bilingual children would decrease over the ages examined in this analysis but the data showed essentially no change. Consistent with this pattern, some studies have reported higher vocabulary scores for monolinguals even in adulthood (e.g., Bialystok, Craik & Luk, 2008; Portocarrero, Burright & Donovick, 2007) so further research is needed to investigate vocabulary growth in bilinguals across the lifespan.

To determine whether the relation between the two languages was a factor in the difference in vocabulary scores, we compared bilingual children whose non-English language was either East Asian or non-Asian. Although this distinction is not linguistically motivated, it clusters around general factors of linguistic and cultural similarity. There was no difference between these subgroups of bilinguals, making bilingualism *per se* the most likely reason for the vocabulary difference.

The vocabulary difference was largely confined to words that are part of home life – a reasonable result given that English is not used as extensively in bilingual homes as it is in those of monolinguals. Importantly, school vocabulary for children in the two groups was more comparable. Thus, bilingual children are not typically disadvantaged in academic and literacy achievement (e.g., Bialystok et al., 2005) or academic uses of spoken language (Peets & Bialystok, 2009) because the linguistic basis of those activities is well established. In this sense, the smaller vocabulary for bilingual children in each language is not an overall disadvantage but rather an empirical description that needs to be taken into account in research designs, especially in tasks that involve verbal ability or lexical processing.

Moreover, the vocabulary deficit for home words in English in the bilingual children is almost certainly filled by knowledge of those words in the non-English language, making it likely that the total vocabulary for bilingual children is in fact greater than that of monolinguals. We consider the results of this item analysis to be preliminary; further research using more detailed categories with a larger sample is required to support our interpretation.

There is considerable evidence demonstrating that adult bilinguals have slower reaction times to name pictures than comparable monolinguals (Gollan, Montoya, Fennema-Notestine & Morris, 2005), even when the naming is carried out in their first and dominant language (Ivanova & Costa, 2008). Two explanations for that effect are weaker connections between words and concepts for bilinguals (Gollan, Montoya, Cera & Sandoval, 2008) and conflict between words in the two languages (Green, 1998). The results of the present study demonstrate another way in which linguistic processing is compromised for bilinguals, this time in childhood. However, the present data do not enable us to determine whether the lower receptive vocabulary in bilingual children and the slower lexical access in bilingual adults reflect a similar underlying mechanism or completely different processes. Further research is necessary to investigate that question.

It is important to establish that bilingual children know fewer words in English than do comparable monolingual speakers of English, especially when all the children are being educated through English in school. This difference, however, does not change the normal properties of their lexical knowledge nor does it interfere with the verbal skills being developed for academic achievement. Bilingual children are constructing the world through two telescopes, and their two vocabularies provide the lenses.

References

Adams, MJ. Beginning to read: Thinking and learning about print. Cambridge, MA: MIT Press; 1990.

Bialystok E, Craik FIM, Luk G. Cognitive control and lexical access in younger and older bilinguals. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2008; 34:859–873.

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

Bracken BA, Murray AM. Stability and predictive validity of the PPVT-R over an eleven month interval. Educational & Psychological Research. 1984; 4:41–44.

Curran PJ, Hussong AM. Integrative data analysis: The simultaneous analysis of multiple data sets. Psychological Methods. 2009; 14:81–100. [PubMed: 19485623]

Dunn, LM.; Dunn, LM. Peabody Picture Vocabulary Test – Third Edition. Bloomington, MN: Pearson Assessments; 1997.

Gollan TH, Montoya RI, Cera C, Sandoval TC. More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. Journal of Memory and Language. 2008; 58:787–814. [PubMed: 19343088]

Gollan TH, Montoya RI, Fennema-Notestine C, Morris SK. Bilingualism affects picture naming but not picture classification. Memory & Cognition. 2005; 33:1220–1234. [PubMed: 16532855]

Green DW. Mental control of the bilingual lexico–semantic system. Bilingualism: Language and Cognition. 1998; 1:67–81.

Ivanova I, Costa A. Does bilingualism hamper lexical access in highly-proficient bilinguals? Acta Psychologica. 2008; 127:277–288. [PubMed: 17662226]

Jastek, JF.; Wilkinson, G. The wide-range achievement test – revised. Wilmington, DE: Jastak Associates; 1984.

Kastner JW, May W, Hildman L. Relationship between language skills and academic achievement in first grade. Perceptual and Motor Skills. 2001; 92:381–390. [PubMed: 11361297]

Lewis, C.; Koyasu, M.; Oh, S.; Ogawa, A.; Short, B.; Huang, Z. Culture, executive function, and social understanding. In: Kewis, C.; Carpendale, JIM., editors. Social interaction and the development of executive function: Special issue of New Directions in Child and Adolescent Development. Vol. 123. 2009. p. 69-85.

Luk G, Bialystok E. Common and distinct cognitive bases for reading in English–Cantonese Bilinguals. Applied Psycholinguistics. 2008; 29:269–289.

Oller DK, Pearson BZ, Cobo-Lewis AB. Profile effects in early bilingual language and literacy. Applied Psycholinguistics. 2007; 28:191–230. [PubMed: 22639477]

Ouellette GP. What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. Journal of Educational Psychology. 2006; 98:554–566.

Peets, KF.; Bialystok, E. Dissociations between academic discourse and language proficiency among bilingual kindergarteners. Poster presented at the Society for Research in Child Development; Denver, CO: 2009.

Portocarrero JS, Burright RG, Donovick PJ. Vocabulary and verbal fluency of bilingual and monolingual college students. Archives of Clinical Neuropsychology. 2007; 22:415–422. [PubMed: 17336036]

Ricketts J, Nation K, Bishop DVM. Vocabulary is important for some, but not all reading skills. Scientific Studies of Reading. 2007; 11:235–257.

Rohde TE, Thompson LA. Predicting academic achievement with cognitive ability. Intelligence. 2007; 35:83–92.

Smith TC, Smith BL, Dobbs K. Relationship between the Peabody Picture Vocabulary Test – revised, wide range achievement test – revised, and Wechsler intelligence scale for children – revised. Journal of School Psychology. 1991; 29:53–56.

Swanson HL, Rosston K, Gerber M, Solari E. Influence of oral language and phonological processing on children's bilingual reading. Journal of School Psychology. 2008; 46:413–429. [PubMed: 19083366]
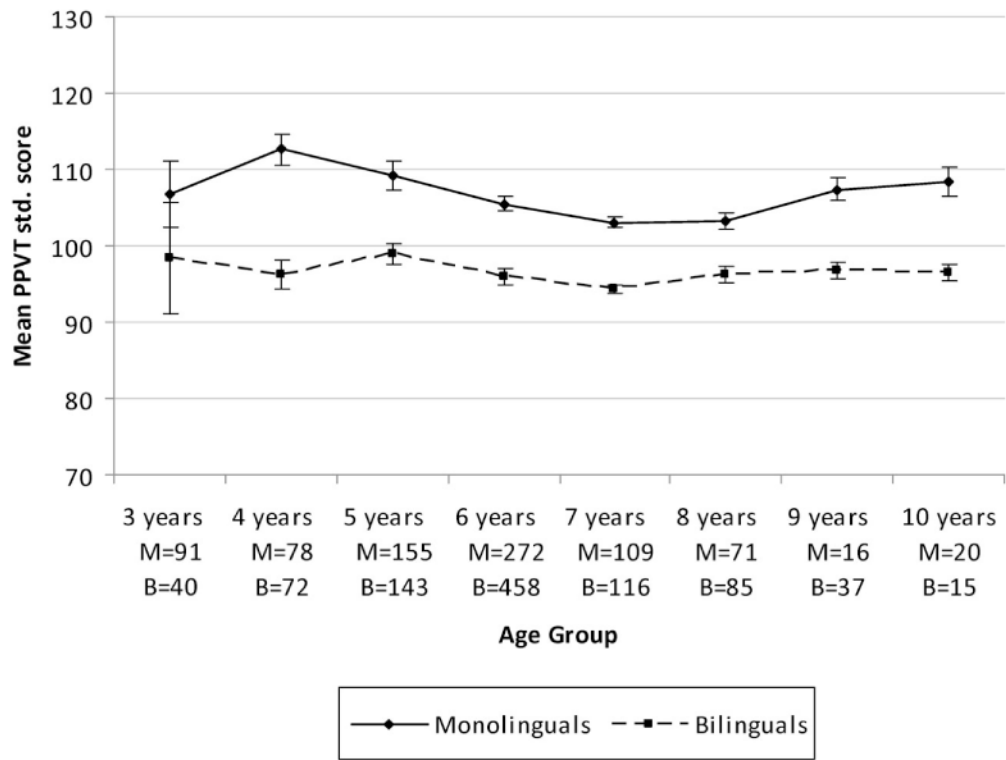
Figure 1.
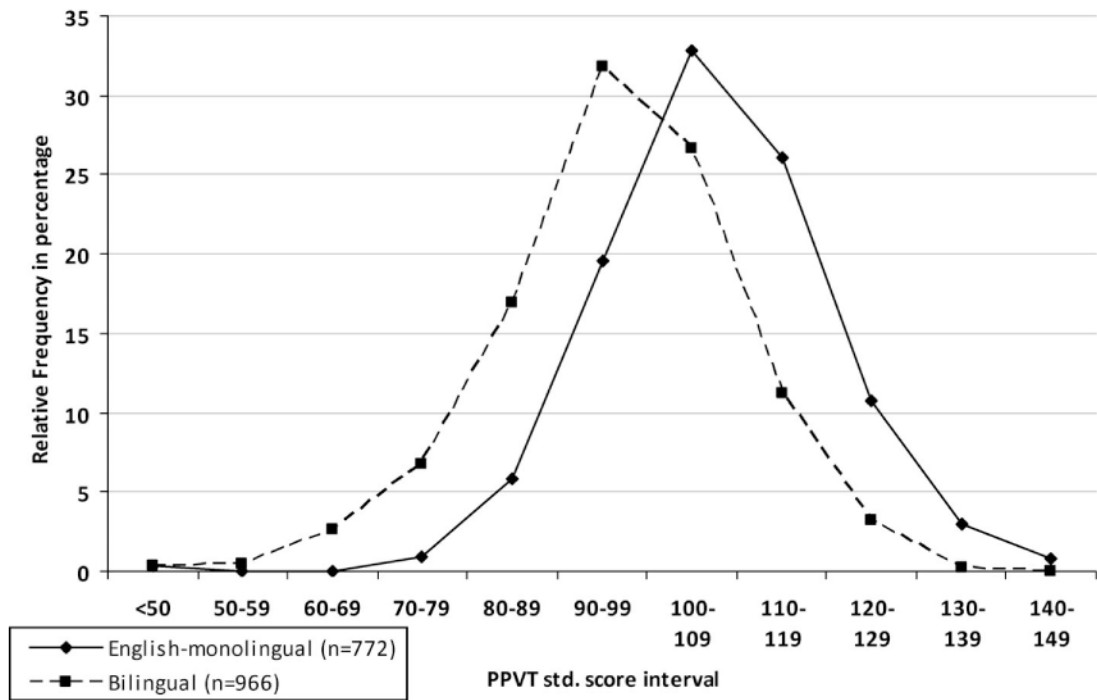Mean PPVT standard score and standard error by age and language group.

Figure 2.
Distribution of PPVT standard scores in English for monolingual (n = 772) and bilingual (n = 966) children.
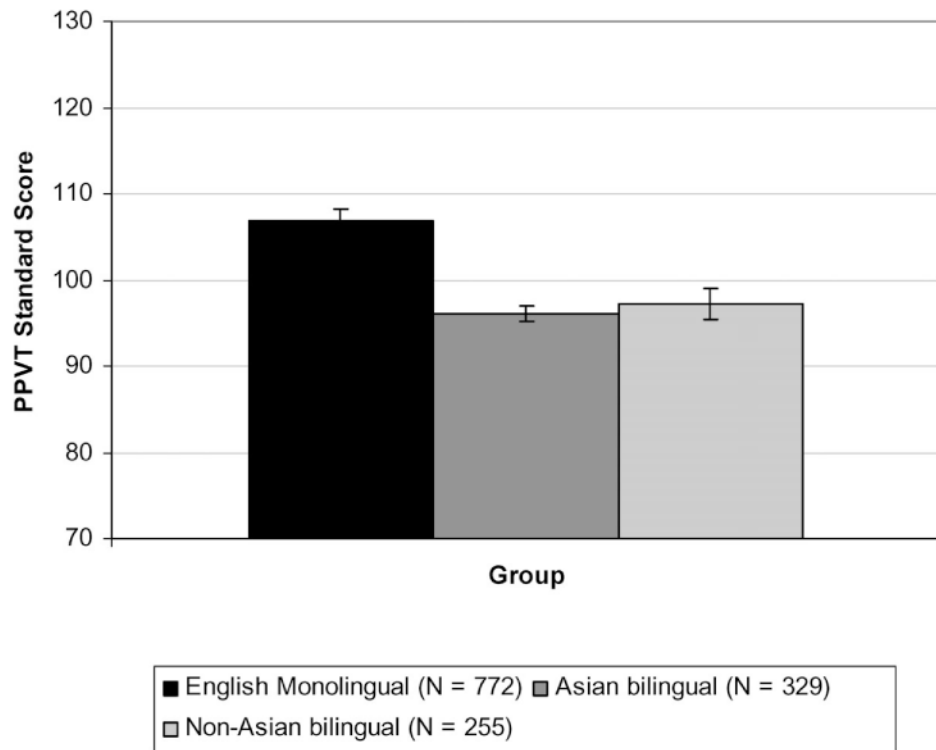
Figure 3.
Mean PPVT standard score and 95% confidence intervals for monolingual English speakers (n=772), bilingual speakers of English and an East Asian language (n = 329), and bilingual speakers of English and a non-Asian language (n = 247).

Table 1

Mean PPVT standard scores and context (home versus school) scores for monolingual and bilingual six year olds

	Monolinguals (n=75)		Bilinguals (n=87)	
	M	**SD**	**M**	**SD**
Age in months	78.0	3.1	77.6	3.0
PPVT standard score	101.8	10.3	96.5	11.9
Home words (percent correct)	77.1	11.3	70.4	12.7
School words (percent correct)	77.7	6.2	75.6	7.5