Bilingualism and Executive Function: Emerging Bilinguals in Head Start

Diana Alicia Santoyo*

Mentor: Dr. Yuuko Uchikoshi
School of Education at UC Davis

Abstract

The aim of this study is to examine dual language learners' (DLLs) bilingual proficiency and the association between proficiency and executive function (EF) components of inhibitory control and cognitive flexibility. Data were collected from 31 Mexican-American and 53 Chinese-American preschoolers enrolled in Head Start centers in Northern California. Results show that, on average, Spanish-English and Cantonese-English DLLs score similarly low in both English and their home language. There were no significant differences between the Mexican-American and Chinese-American children on the oral language measures in both L1 and English. Furthermore, despite their low oral proficiency, EF scores suggest that on average they score similarly to their monolingual-English speaking peers. DLLs with stronger language skills in both languages tended to do better on EF than those with low skills in both languages. Implications for educational policy and classroom instruction will be discussed.

Theoretical Framework

In the last decade, dual language learners (DLLs; children who are exposed to and learning through two languages) in the United States have increased from 20 million to 23 million (or from 28% to 32% of all children). DLLs are more likely to live under poverty and have less-educated parents compared to their monolingual English-speaking peers (Child Trends, 2016). Although the majority of DLLs are from Spanish-speaking households, most recently, there has been a dramatic increase in DLLs from Chinese-speaking households (Child Trends, 2016). Much of the past studies have been with Spanish-English DLLs; not much research has been conducted with DLLs with other language combinations, such as Chinese-English DLLs (Hammer et al., 2014). In particular, how similar or different Chinese-English and Spanish-English DLLs in Head Start programs are under-explored.

Moreover, despite research suggesting that bilingualism may be associated with some cognitive benefits, DLLs from low-income families lag significantly behind their monolingual peers on school readiness measures at kindergarten entry, and this achievement gap widens with age. Furthermore, current literature has contrasting views on whether or not bilingualism promotes children's school readiness skills, cognitive ability and socioemotional outcomes. Some researchers have found no bilingual advantage in EF in inhibitory control or set shifting (Paap, Johnson, & Sawi, 2015). Yet, others suggest that bilingualism enhances children's executive function due to their ability to inhibit one language while speaking another language (Bialystok, 2001). This is known as the bilingual advantage in executive control (Bialystok, 2001). The effects of bilingual advantage appear to be specific in some executive domains, such as inhibitory control (Bialystok & Senman, 2004; Bialystok & Shapero, 2005) and cognitive flexibility (Costa, Hernández, & Sebastián-Gallés, 2008; Prior & Macwhinney, 2010). Essentially, it is important to look into EF because it is responsible for the tasks we complete daily. The mental processes enable our capacity to plan, organize, and monitor the execution of behaviors. The development of EF throughout childhood and onward is the influencer to behavior, mental flexibility, and self-control.

Thus, more research is needed to understand (1) the bilingual proficiency of both Spanish-English and Chinese-English DLLs, and (2) the relationships between young DLLs' bilingual proficiency and EF skills.

Research Question

The aims of this study are to examine (1) the levels of bilingual proficiency in young Chinese-American and Spanish-American children and (2) the association between language proficiency in the dominant and nondominant languages of Spanish and English and Cantonese and English and the executive function (EF) components of inhibitory control and cognitive flexibility with emerging Spanish-English and Cantonese-English bilingual preschoolers enrolled in Head Start programs in Northern California.

Methods

Participants

A total of 84 preschoolers attending Head Start programs and their parents in Northern California were recruited for this study. In total, there were 31 Spanish-speaking Mexican Americans [MA] and 53 Cantonese-speaking Chinese Americans [CA] DLLs. All children were between 36 and 60 months of age. The average age of the participants was 49.23 months. There were no differences in the average age of the MA and CA participants. Because of our focus on low-income families, the children and parents were recruited from Head Start centers in Northern California. With respect to language and literacy skills, the children were exposed to Spanish or Cantonese at home and had to be able to produce two-word telegraphic speech in Spanish or Cantonese before age three (based on index used in other studies of DLLs). Data collection by trained research assistants speaking either native English, native Spanish or native Cantonese was done in the children's homes over a 3-hour period

Parents of the children have the following demographics. On average, they have 11 years of education with a per capita income of $8,571. In addition, these parents have been residing in the United States on average of 9 years. There does not seem to be a significant difference between MA and CA families on parental education, socioeconomic status, and residency in the U.S.

Measures

Language Proficiency

Language proficiency was measured with the Woodcock-Johnson IV Tests of Oral Language (WJ IV OL; Schrank, Mather, & McGrew, 2014). Specifically, we used the Picture Vocabulary, Oral Comprehension, and Understanding Directions subtests in both the first language (L1; Spanish or Cantonese) and English. The Picture Vocabulary subtest (WJPV) requires the child to name both familiar and unfamiliar pictures, ordered by increasing difficulty, with each response scored as correct or incorrect by the assessor. The Oral Comprehension subtest (WJOC) requires the child to listen to and supply a missing word to the end of a sentence, or related group of sentences. The Understanding Directions subtest (WJUD) requires the child to look at a picture and respond to verbal requests to point to certain items on the picture. The English version was used to assess English oral skills. The Spanish version was used to assess Spanish oral skills. The Spanish version was translated to Cantonese and used to assess

Cantonese oral skills as done in past research (e.g. Uchikoshi, 2013).

Executive Function

EF was measured with the NIH ToolBox Dimensional Change Card Sort Test and Flanker Test and the Head-Toes-Knees-Shoulders Task.

NIH ToolBox Dimensional Change Card Sort Test (Age 3-7) (DCCS)

DCCS is a measure of cognitive flexibility -- the ability to shift attention between tasks. This test has been widely used to measure executive function in children (Zelazo et al., 2013). During the task, participants are presented two targeted cards on an iPad, and then assessed on their ability to distinguish the dimensional differences (shape or color) by selecting a series of test cards. This test has shown high test-retest reliability (ICCs=.86–.95, Zelazo et al., 2013).

NIH ToolBox Flanker Test

Flanker test is an inhibitory control and attention measures. This test was originated from the Eriksen flanker task (Eriksen & Eriksen, 1974). In the flanker test, participants are presented with 5 arrows horizontally. They are asked to indicate where the middle arrow is pointing, while inhibiting attention to the arrows next to it. The NIH Toolbox first presented fish instead of arrows, this is designed to get children's attention and engage them more with the task. Difficulty increases as the child continues and eventually, trial will no longer be presented with fish but arrows. The flanking stimuli can sometimes be congruent with the middle orientation and sometimes can be incongruent. This test has shown high test-retest reliability (ICCs=.86–.95, Zelazo et al., 2013).

Heads-Toes-Knees-Shoulders (HTKS)

In this task, children are asked to play a game in which they must do the opposite of what the experimenter says. The HTKS task has been conceptualized by Ponitz et al., (2008) as a measure of inhibitory control, working memory and attention focusing.

Statistical Analysis

Group differences between MA and CA were examined with t-tests using the Satterthwaite approximation. The means, standard deviation, and ranges of the scores are presented in Table 1. Groups were collapsed and grouped using cluster analysis and results of the cluster analysis were used to examine relationships with EF.

Results

Descriptive Analysis

The performance of MA and CA children in language and EF measures is summarized in Table 1.

Language Performance

As shown in Table 1, although the range in scores was large, on average, children scored over one standard deviation below the monolingual age-equivalent norms on the English oral proficiency measures. On the picture vocabulary task, there were no differences in English expressive vocabulary standard scores between the two groups; $t(49.42)=-.34$, $p=.7357$. Similarly on the oral comprehension task, there were no differences in standard scores between the two groups; $t(46.01)=-1.19$, $p=.2405$. This was also true for the understanding directions task, where there were no differences in the standard scores between the two groups; $t(64.1)=-.47$, $p=.6364$.

Table 1. *Descriptive statistics of language measures for all children (n = 84), all Mexican-American children (n = 31), all Chinese-American children (n = 53).*

Variables	Total Range	Total Mean (SD)	MA Children (SD)	CA Children (SD)
Eng. WJPV Range	0-24	12.06 (5.17)	12.23 (5.53)	11.96 (4.99)
Eng. WJPV SS	40-117	76.88 (16.88)	77.79 (18.73)	76.35 (15.88)
Eng. WJOC Raw	0-12	1.15 (2.14)	1.32 (1.89)	1.04 (2.29)
Eng. WJOC SS	48-118	68.95 (14.56)	71.71 (16.78)	67.33 (13.01)
Eng. WJUD Raw	0-21	7.95 (5.58)	8.40 (5.56)	7.67 (5.64)
Eng. WJUD SS	40-125	80.16 (16.93)	81.32 (15.28)	79.48 (17.95)
L1 WJPV Raw	0-22	9.91 (6.45)	10.90 (7.01)	9.31 (6.08)
Sp. WJPV SS	40-110	69.14 (24.91)	69.14 (24.91)	- -
L1 WJOC Raw	0-8	1.30 (1.93)	1.89 (2.41)	0.96 (1.49)
Sp. WJOC SS	40-94	61.07 (18.55)	61.07 (18.55)	- -
L1 WJUD Raw	0-21	4.97 (4.20)	5.37 (5.03)	4.74 (3.63)
Sp. WJUD SS	40-103	64.75 (19.84)	64.75 (19.84)	- -

Note. SS = Standardized scores. L1 includes Spanish (Sp) and Cantonese (Ca).

Of the three English oral proficiency measures, when compared to the published age-matched monolingual English population norms, on average, children did best on the understanding directions subtest. This may be due to the fact that this is a receptive measure and the child only needs to point to the answers, as opposed to the other measures that require the child to verbally respond. For the other two tasks, the children, on average, scored more than 1.5 standard deviations lower than the published age-matched English monolingual norms.

When examining the Spanish standard scores, we found that on average the children scored over one standard deviation below the monolingual age-equivalent norms. The raw scores in Spanish and Cantonese inform us that on average both groups scored similarly low on their home language measures. T-test analysis using raw scores show no differences between the MA and CA groups (L1 picture vocabulary task: t(54.82)=-1.03, p=.3072; L1 oral comprehension task: t(42.33)=-1.70, p=.0958; L1 understanding directions task: t(47.47)=-.60, p=.5521).

Executive Function Performance

The means, standard deviation, and ranges of executive function performance are presented in Table 2.

Table 2. *Descriptive statistics of executive function measures for all children (n = 84), all Mexican-American children (n = 31), all Chinese-American children (n = 53).*

Variables	Total Range	Total Mean (SD)	MA Children (SD)	CA Children (SD)
NIH Raw	0-40	12.66	13.75	10.48
(Flanker)		(9.39)	(9.02)	
(9.40)				
NIH Age-Corrected	0-122	80.13	84.96	72.24
(Flanker)		(35.48)	(34.40)	(43.00)
NIH Raw	0-36	7.81	5.63	8.00
(DCCS)		(9.60)	(8.62)	(11.01)
NIH Age-Corrected	0-115	77.37	56.29	87.66
(DCCS)		(36.70)	(46.53)	(26.49)
HTKS Raw	0-47	4.83	4.00	5.31
		(9.96)	(8.23)	(10.89)

When compared to monolingual age-equivalent norms on the EF tasks, on average, the children scored over 1.5 standard deviations below the published mean for both the Flanker and DCCS measures. However, the variation in scores were large. Interestingly, the MA group scored higher than the CA group on the Flanker task, t(75.84)=-2.03, p=.0454, but there were no significant group differences on the DCCS task, t(44.02)=1.94, p=.0589, and the Head-Toes-Knees-Shoulders Task, t(68.97)=.59, p=.5548.

Cluster Analysis

As there were no significant differences between the English oral language scores and the L1 oral language scores, both MA and CA groups were merged to see if there were any differences among the combined abilities of L1 and English among the participants. Figure 1 shows the relationship between L1 picture vocabulary and English picture vocabulary by group. Figure 1 suggests that both MA and CA participants had similar varied English and L1 vocabulary knowledge.

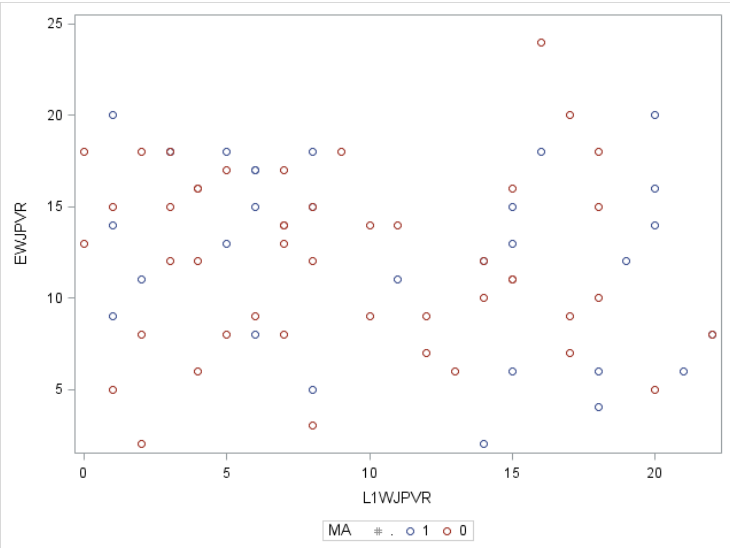

Figure 1. Relationship between English picture vocabulary (EWJPVR) and L1 picture vocabulary (L1WJPVR) by Mexican American (blue) and Chinese American (red) group.

Using agglomerative cluster analysis (Ward's method) on all subjects with raw scores on the L1 and English picture vocabulary subtest, four clusters emerged. Examining the four clusters revealed that the largest cluster included children who had higher English vocabulary scores and low L1 scores (English dominant), the next largest had children with high L1 scores and low English scores (L1 dominant), followed by those who had low L1 and low English (low bilingual), and those with equivalently high English and L1 scores (ideal bilingual). The means, standard deviation, and ranges of the scores for vocabulary, and EF subtests by clusters are presented in Table 3.

Table 3. *Four clusters analysis.*

Variables	Ideal Bilingual (n=10)	English Dominant (n=31)	L1 Dominant (n=23)	Low Bilingual (n=12)
Eng. vocab.	17.60	15.61	8.43	6.83
Mean (SD)	(3.06)	(2.22)	(2.90)	(2.66)
L1 vocab.	17.50	5.00	15.91	4.33
Mean (SD)	(2.01)	(2.95)	(3.42)	(2.67)
NIH Age-Corrected	96.90	84.07	92.26	59.00
(Flanker) Mean (SD)	(9.71)	(27.10)	(25.25)	(52.88)
NIH Age-Corrected	93.80	76.00	75.61	69.67
(DCCS) Mean (SD)	(5.96)	(35.02)	(41.92)	(42.80)
HTKS Raw	1.80	4.50	7.39	0.54
Mean (SD)	(4.69)	(10.11)	(11.78)	(1.29)

Executive Function

Results show a large variation on the NIH ToolBox EF tasks among the groups. Interestingly, the variation on the NIH ToolBox EF measures was the smallest for the "ideal bilingual" group.

A one-way between subjects ANOVA was conducted to compare the means of executive function tasks by clusters. There were significant differences on the Flanker task among the four cluster groups [$F(3, 72) = 3.11$, $p = .0314$]. Post hoc comparisons using the Tukey HSD test indicated that the mean score for the ideal bilingual group ($M = 96.9$, $SD = 9.71$) was significantly different than the low bilingual group ($M = 59$, $SD = 52.88$). However, the other groups did not significantly differ from each other.

Although the means varied among the four groups, a one-way between subjects ANOVA revealed no significant differences on the DCCS task among the four cluster groups [$F(3, 72) = .88$ $p = .4560$]. Similarly, the HTKS task showed large variation in the means among the four groups, and a one-way between subjects ANOVA revealed no significant differences on the HTKS task among the four cluster groups [$F(3, 69) = 1.65$ $p = .1854$].

Discussion

The findings of this study suggest that at the beginning of Head Start, Spanish-English DLLs and Cantonese-English DLLs may have more similarities than differences in the language and executive function skills. On average, both sets of DLLs appear to be struggling to increase their oral proficiency in both languages. In English, on average, both groups were over 1.5 standard deviations below the published English norm. Their L1s were on average similar to their English oral proficiency. This supports past studies where DLLs performed below the monolingual English children on oral proficiency (Hammer et al., 2014). At the same time, there appears to be large variations in proficiency with some DLLs achieving bilingual proficiency, while many were language dominant in either English or their home language. This may be due to a variety of factors as shown in previous research, such as language exposure and mother's education (Hammer et al., 2014).

As for executive function, similar to past research, there appears to be differences in results based on the task. Group differences were seen in the Flanker task, yet all groups performed similarly on the DCCS and HTKS tasks. Flanker task tests for inhibition. Our results suggest that children who have stronger language skills in both languages tend to do better with inhibitory control when compared to children who have weaker language skills in both languages. This is in line with previous research where bilingual children performed better in the inhibitory control task than monolinguals (Bialystok & Senman, 2004; Bialystok & Shapero, 2005).

Implications

One implication is that is all DLLs need to increase their oral proficiency. They need to be exposed to a variety of vocabulary and have more opportunities to use their language in either or both of their languages. Past research has shown that the quality and quantity of their language exposure has an impact on children's language development (Rowe, 2012).

Another implication is that children can improve their oral proficiency in their home language without hindering their English language development. A group of DLLs were able to reach proficiency in both of their languages. It appears that if children are able to achieve some level of bilingual proficiency, this may impact their EF systems.

Limitations

The limitations of this study involve the sample size and criteria. A larger sample would give more power to this study. Since the participants came from a specific area, namely Northern California, and were from a specific population (MA and CA), these results are not generalizable to everyone who speaks more than one language. Another limitation concerns the hand dominance of the children when doing the NIH ToolBox. Children tended to forget to put their dominant hand on home base or began using their nondominant hand. This can cause a discrepancy in time calculation and scoring.

Significance

The results from this study show that the Chinese-American and Mexican-American children in this study may have more similarities than differences. This may have implications for preschool and head start classrooms. All teachers for all students need to focus on improving the children's oral proficiency, especially vocabulary, in either or both of the languages of bilingual children. Parents can also be advised to interact with their children in their home language as children proficient in both languages may have an advantage in particular areas of EF, such as inhibition. As this is one of the few studies examining young DLLs, more research is needed to further examine the relationship between bilingualism and executive function at younger ages.

References

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. New York, NY, US: Cambridge University Press http://dx.doi.org/10.1017/CBO9780511605963

Bialystok, E. & Senman, L. (2004). Executive processes in appearance–reality tasks: The role of inhibition of attenRtion and symbolic representation. *Child Development*, *75*, 562-579. doi:10.1111/j.1467-8624.2004.00693.x

Bialystok, E. & Shapero, D. (2005). Ambiguous benefits: the effect of bilingualism on reversing ambiguous figures. *Developmental Science, 8*, 595-604. doi:10.1111/j.1467-7687.2005.00451.x

Child Trends. (2016). Dual language learners. Retrieved from https://www.childtrends.org/indicators/dual-language-learners

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition. 106*(1), 59-86. https://doi.org/10.1016/j.cognition.2006.12.013.

Eriksen, B. A. & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. https://doi.org/10.1186/1744-9081-3-42

Hammer, C. S., Hoff, E., Uchikoshi, Y., Gillanders, C., Castro, D. C., & Sandilos, L.E. (2014).
The language and literacy development of young dual language learners: a critical review. *Early Childhood Research Quarterly, 29*, 715-733. doi: 10.1016/j.ecresq.2014.05.008

Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 69,* 265-278. http://dx.doi.org/10.1016/j.cortex.2015.04.014

Ponitz, C. C., McClelland, M. M., Matthews, J. S., & Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology*, *45*(3), 605–619. https://doi.org/10.1037/a0015365

Prior, A., & Macwhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and Cognition*, *13*(2), 253–262. https://doi.org/10.1017/S1366728909990526

Rowe, M. L. (2012). A Longitudinal Investigation of the Role of Quantity and Quality of Child-Directed Speech in Vocabulary Development. *Child Development, 83(5),* 1762-1774.

Schrank, F. A., Mather, N., & McGrew, K. S. (2014). *Woodcock-Johnson IV Tests of Oral Language*. Rolling Meadows, IL: Riverside.

Uchikoshi, Y. (2013). Predictors of English reading comprehension: Cantonese-speaking English language learners in the U.S. *Reading and Writing: An Interdisciplinary Journal, 26*(6), 913-939. http://dx.doi.org/10.1007/s11145-012-9398-z

Zelazo, P. D., Anderson, J. E., Richler, J. , Wallner-Allen, K. , Beaumont, J. L., & Weintraub, S. (2013). II. NIH ToolBox Cognition Battery (CB): Measuring executive function and attention. *Monographs Society Res Child*, *78*, 16-33. doi:10.1111/mono