CHARACTERIZING BILINGUAL TODDLERS' VOCABULARY

by

Kimberly Leon

A Thesis Submitted to the Faculty of the

DEPARTMENT OF PSYCHOLOGY

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF ARTS

In the Graduate College

THE UNIVERSITY OF ARIZONA

2022

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Master's Committee, we certify that we have read the thesis prepared by: Kimberly Leon
titled:

and recommend that it be accepted as fulfilling the thesis requirement for the Master's Degree.

_______________________________ Date: Aug 13, 2022
Rebecca Gomez

_______________________________ Date: Aug 13, 2022
Janet Nicol

_______________________________ Date: Aug 13, 2022
Louann Gerken

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to the Graduate College.

I hereby certify that I have read this thesis prepared under my direction and recommend that it be accepted as fulfilling the Master's requirement.

_______________________________ Date: Aug 13, 2022
Rebecca Gomez
Thesis Committee Chair
Psychology

ACKNOWLEDGEMENT

First, I would like to thank my advisor, Dr. Rebecca Gomez. Thank you for your continuous support and encouragement. Additionally, I would like to thank my committee members, Dr. LouAnn Gerken and Dr. Janet Nicol, for their advice and support. Lastly, I would like to express my gratitude to those who supported me during the completion of this thesis.

LAND ACKNOWLEDGEMENT

We respectfully acknowledge the University of Arizona is on the land and territories of Indigenous peoples. Today, Arizona is home to 22 federally recognized tribes, with Tucson being home to the O'odham and the Yaqui. Committed to diversity and inclusion, the University strives to build sustainable relationships with sovereign Native Nations and Indigenous communities through education offerings, partnerships, and community service.

Table of Contents

List of Tables

List of Figures

Abstract

Bilingual children's vocabulary remains difficult to assess due to the multiple factors that must be considered and lack of norms typically available. One issue in characterizing bilingual vocabulary is how to count translation equivalents (TEs), which refer to labels in each language that refer to the same concept. This study investigated a new potential vocabulary measure that modifies how much weight to give to TEs when counting vocabulary depending on the child's age. We extended this new measure with Spanish-English bilingual toddlers and compare the weights with the original dataset with which the measure was initially developed. Looking at the vocabulary scores, our sample of monolingual and bilingual children did not differ on the new measure. Although there were no statistically significant differences in vocabulary, there were subtle numerical differences at certain age bands. When analyzing the scores further, higher weights were given to older bilingual children. Given what is known about bilingual language development, we probed the measure's theoretical assumptions. Considering the divisive evidence for whether learning TEs creates different challenges at different ages, we suggest that more work is required for the development of future measures vocabulary that aim to predict, characterize, or assess bilingual vocabulary.

Introduction

Vocabulary is a useful tool to measure language ability and identify potential language delays or impairments (Dale, 1991; Ellis & Thal, 2008). It is a known predictor of language, literacy skills and academic achievement (Duff et al., 2015; Marchman & Fernald, 2008). Thus, measuring vocabulary correctly is imperative. Importantly, past research has lacked the ability to accurately measure vocabulary in bilingual populations (Gonzalez-Barrero et al, *2020*).

Bilingual children are at higher risk for being under or over diagnosed with a language delay or problem despite research showing equal likelihood compared to monolingual children (Crutchley, 2000; Kohnert, 2010; Paradis, Genesee & Crago, 2010). The historically negative research on bilingualism as well existing research methodology contribute to these issues. Current vocabulary measures mostly depict bilingual children as significantly different than monolingual children. This stems from the way these measures are constructed, from a monolingual perspective. The existing measures struggle to capture unique aspects of bilingual vocabulary, especially translation equivalents. Translation equivalents (TEs) are labels in different languages that refer to the same concept (Core et al., 2013; Marchman & Martinez-Sussmann, 2002). Children who are bilingual will inherently have more than one label for the same concept. As noted, there is no consensus on how to count these.

The results tend to be mixed based on the various measures of vocabulary. Total vocabulary is calculated as the sum of words known in both languages. This measure may appear to overcount and does not provide comparable scores to monolingual children. Most bilingual children will appear to have higher or at times equal vocabulary to monolingual peers using this measure (Bosch & Ramon-Casas, 2014; Core et al., 2013, Hoff et al., 2012, De Houwer et al., 2014; Junker & Stockman, 2002; Pearson et al., 1993, Thordardottir et al., 2006). Dominant vocabulary is calculated as the sum of words known only in the dominant language. Non-dominant vocabulary is calculated as the sum of words known only in the non-dominant language. Both measures discard information regarding a child's ability as they only consider one of a child's two languages. Unsurprisingly, dominant and non-dominant vocabulary measures tend to depict bilingual vocabulary as significantly smaller compared to monolinguals (Bialystok et al., 2010, Cattani et al., 2014). Lastly, concept vocabulary refers to the sum of concepts uniquely known in language one, concepts uniquely known in language two and TEs (only counted once). Bilingual children also appear to have significantly lower vocabulary using concept vocabulary (Bosch & Ramon-Casas, 2014; Core et al., 2013, Hoff et al., 2012, De Houwer et al., 2014; Junker & Stockman, 2002; Pearson et al., 1993, Thordardottir et al., 2006). Therefore, assessing vocabulary in bilingual children remains a challenging task for researchers and clinicians due to inconsistent results.

Importantly, bilinguals are not two monolinguals in one (Grosjean, 1989). Bilingual researchers continue to argue the importance of including both languages in assessment (Nicoladis & Genesee, 1997). Hence, multiple assessments have been adapted into a bilingual version to better fit that population (Arizmendi & Alt, in prep; Nicoladis & Wiebe, 2019). Thus, adapting current vocabulary measures to better fit the bilingual population is a clear next step.

Norms must be readily available for the target population to benefit from any specific measure, but norms are rarely available. Truthfully, developing norms is a costly procedure and more difficult for bilingual children given the variability in language pairs, the amount of exposure to each language, and other factors (Gonzalez-Barrero et al., 2020). However, the necessity of norms is evident. Having comparable norms to monolingual children can help provide insight into the similarities and differences in bilingual children's respective language development (Byers-Heinlein et al., 2013; Gonzalez-Barrero et al., 2020). Most importantly, comparable norms allow for a quick assessment of language development and can hint at the need for a more thorough clinical assessment or services (Bedore & Pena, 2008, Gonzalez-Barrero et al., 2020). Furthermore, adopting a measure that allows us to compare both groups would increase the scope of existing research by allowing for more inclusive and generalizable research by accurately charactering vocabulary in bilingual children.

Given what we know about bilingualism, an adaptation may be an adequate transformation of the existing vocabulary measures. Bilingualism is highly common and a growing experience. Around 21% of school age children in the US spoke another language than English at home in 2017 (U.S. Census Bureau, 2017). Bilingual infants achieve milestones similarly to monolingual infants (Werker & Byers-Heinlein, 2008). Despite persistent myths, bilingual children can differentiate between their languages and learn in both (Byers-Heinlein & Lew-Williams, 2013). As seen, many bilingual adults exist and are highly capable in their languages.

Homing in on language development, bilingual and monolingual children share predictable developmental patterns (Werker & Byers-Heinlein, 2008). The noun bias has been long demonstrated in English speaking monolingual children. Recent work with bilingual toddlers aligns with patterns observed during play between monolingual mother-child dyads. Bilingual toddlers learning both a noun-friendly (English) and a verb-friendly language (Mandarin) exhibited noun dominance in their expressive vocabularies with a more prevalent noun dominance in English (Setoh et al., 2021). The shape bias is another lexical bias important to early language development. Spanish-English bilingual toddlers show a shape bias in novel noun generalization tasks at 21 months when tested in English but not in Spanish (Schonberg, Russell & Luna, 2019). This work points to similarities across early language development by language in both bilingual and monolingual groups.

An obvious difference between bilinguals and monolinguals is language experience. Bilinguals will have less exposure and less experience with each of their languages compared to a monolingual. This reality stems from time constraints and is an important consideration when measuring vocabulary. Vocabulary may also be largely impacted by the contexts in which children are learning language (Bialystok et al., 2010; Hoff et al., 2010). The amount of vocabulary further varies according to the amount of language exposure (Hoff et al., 2014). Additionally, language use is a contributing factor, including how children themselves use their languages (Hoff, 2020).

Past research delineates how bilingual and monolingual children rely on the same word learning mechanisms (Byers-Heinlein, Burns & Werker, 2010; Byers-Heinlein et al., 2013; Curtin et al., 2011; Weatherhead et al., 2021). Soon after birth, monolingual and bilingual infants equally discriminate between English and Tagalog suggesting the same perceptual mechanism supports early language discrimination (Byers-Heinlein, Burns & Werker, 2010). Additional evidence comes from, PRIMIR, a theoretical framework focusing on the bidirectional relationship between speech perception and vocabulary (Curtin et al., 2011). One of its main premises is that bilingual and monolingual children are supported by the same learning mechanisms. The switch task, where children learn a label-object mapping and are presented with the same vs. different mapping, has been studied in the context of associative word learning. (Byers-Heinlein et al., 2013). Performance on this task was similar in bilinguals and monolinguals with 14-month-old infants from both language backgrounds accurately detecting the different mapping yet failing to detect it at 12 months of age. This suggests associative word learning is a robust strategy that supports language development despite language background. Such data may suggest that bilingual children rely on the same word learning mechanisms.

Mutual exclusivity (ME) is one word learning heuristic that has been noticeably studied in research with bilingual toddlers. ME refers to the assumption that an object has only one label. ME is arguably intriguing given its inherent violation when learning language as a bilingual child. Past research has consistently found bilingual toddlers to employ ME at significantly lower rates with lower disambiguation and lower retention than their monolingual counterparts (Rocha et al., 2022). However, these differences have been recently brought into question. Weatherhead and colleagues found monolingual children to employ ME similarly to bilingual children, including violating ME, in certain word learning tasks where the context was manipulated by using speakers whose race was familiar or unfamiliar to them (2021). Both groups failed to employ ME when the speaker's race is unfamiliar but applied it when the speaker's race was familiar. Importantly, these findings highlight how monolingual and bilingual children rely on the same word learning mechanisms but may use them differently as a function of their experiences.

With these considerations in mind, finding a way to capture vocabulary for monolingual and bilingual children that allows for equitable comparisons may be appropriate. Gonzalez-Barerro and colleagues (2020) developed a new metric, the Bilingual Adjusted Vocabulary (BAV), focusing on how to modify the way TEs are currently counted. They utilized a large existing dataset of expressive vocabulary data obtained from 184 French-English bilinguals, 112 English monolinguals and 108 French monolinguals, ages 18-33 months using the MacArthur Bates Communicative Development Inventory: Words and Sentences form in English and French.

Their tactic focused on how to count TEs (given 2 credits in total vocabulary and 1 credit in concept vocabulary). They argued learning a TE remains a difficult task, even if it is only learning the label and mapping it on to a semantic representation, especially at younger ages. They adopted a mathematical approach to figure out how much credit TEs should contribute taking age into account (2 credits if counted twice and 1 credit if counted once). They fit a model with total vocabulary the outcome, age the 1st predictor, the number of TEs the 2nd predictor and the 3rd predictor the interaction between age and number of TEs. They hypothesized that a significant interaction would indicate a need to give different credit at different ages. This assumes the association between the number of TEs and total vocabulary depends on age, as observed in a longitudinal analysis observing proportional increase of TEs from 16 to 30 months of age (David & Wei, 2008). They combined the model estimated coefficients for the fixed effect of number of TEs plus its interaction with age to compute adjustment coefficients. These were truncated to 2 as the expected range is 1-2 given that children should have a minimum of one label and a maximum of two labels for the same concept in both languages. Then, they multiplied the number of TEs by the truncated adjustment coefficients and added it to the number of words only known in English plus the number of words only known in French. This resulted in the BAV measure. Using this measure, they fit the same model as with the traditional measures. Although bilinguals differed from monolinguals in predictable ways using the traditional measures, bilinguals and monolinguals showed no significant differences using the BAV measure.

The presence of TEs is observed early in children's vocabulary (De Houwer et al., 2006) and varies with age and according to the languages involved (Davis & Wei, 2008; Gonzalez-Barrero et al., 2020). Given how TEs can vary in relation to language-specific exposure, it is imperative to expand this approach to a different population. Notably, the variation in TEs may be seen between symmetric and asymmetric bilinguals. Prior work posits the balanced exposure in symmetric bilinguals will allows for more opportunities to learn TE thus leading to a higher presence of TEs compared to asymmetric bilinguals (Legacy et al., 2017). Therefore, accounting for language symmetry is key for the successful implementation of this measure. Thus, we expand this approach to a population of Spanish-English bilingual toddlers with Spanish monolinguals and English monolinguals as reference points. Specifically, we broaden the relative amount of language variation acceptable to further test the scope and efficacy of this metric. In turn, this will allow us to differentiate the metric's efficacy in symmetric vs. asymmetric populations. Therefore, the primary aims of this study are to replicate the findings reported in Gonzalez-Barrero et al (2020), extend their approach to a dataset containing Spanish-English bilinguals, and analyze the differences, if any, between samples.

Methods

Participants

Data were obtained from Wordbank, an open repository for researchers to share vocabulary data (Frank et al., 2016). The dataset consists of some repeated measures per child thus we will refer to our sample as the number of administrations. Our data consisted of 175 Spanish-English bilinguals, 289 Spanish monolinguals and 126 English monolinguals, ages 18-33 months (see table 1 below for descriptive statistics). Children were excluded if they were preterm (born before 36 weeks gestation) or had low birth weight (less than 5.5 lbs). Age did not vary significantly according to the distinct groups of Spanish-English bilinguals (M = 25.62 months, *SD* = 4.27), Spanish monolinguals (M = 24.89 months, *SD* = 4.12) and English monolinguals (M = 24.62 months, *SD* = 3.26); F(2, 587) = 2.7361, p = 0.065). Spanish and English monolinguals were combined in Table 1.

Table 1

Descriptive Statistics (Means and SDs)

Variable	Monolingual (*n* = 415)	Bilingual (*n* = 175)
Age in months	24.69 (*3.87*)	25.62 (*4.27*)
Age range	18-33	18-33
% Female	42%	49%
% English Dominant	30%	35%

Note. *n* is based on number of administrations.

Given our aim to test the scope and efficacy of this metric, we utilized a generous exposure criterion. Infants were classified as bilinguals having non-zero exposure to two languages. Notably, the asymmetric language exposure this permits contrasts with more stringent criteria imposed in Gonzalez-Barrero et al. (*2020*). Exposure to each language was measured using the vocabulary survey's basic information form (Fenson et al., 2007; Jackson-Maldonado et al., 1993). Bilingual infants were classified as Spanish or English dominant based on the language with the highest exposure. Around 35% of the bilingual toddlers were English-dominant. Exposure to the dominant language ranged from 44% to 99% in Spanish-dominant toddlers and from 18% to 97% in English-dominant children.

Apparatus

Expressive vocabulary data was available for each child. Data were collected using the MacArthur-Bates Communicative Developmental Inventory: Words & Sentences (MBCDI-WS) for English and the MacArthur Inventarios Del Desarrollo de Habilidades Comunicativas: Palabras & Enunciados (MIDHC-PE) for Spanish (Fenson et al., 2007; Jackson Maldonado et al., 2003). These surveys ask parents to mark off the words their child produces independently and have been shown to be reliable measures (Marchman & Martinez-Sussmann, 2002). TEs were coded by extending the concepts identified by Marchman in the vocabulary surveys for Spanish and English (1999), following the approach used by Gonzalez-Barrero and colleagues (2020).

Procedure

We calculated four measures of vocabulary for each child including total, dominant, non-dominant and concept vocabulary following Gonzalez-Barrero et al. (2020). Total vocabulary was calculated as the total number of words across both languages, dominant vocabulary as the total number of words in the dominant language, non-dominant vocabulary as the total number of words in the non-dominant language, and concept vocabulary as the sum of concepts uniquely known in language one, concepts uniquely known in language two and TEs (only counted once).

We calculated the BAV measure in multiple steps. We first centered age (continuous) at 25 months, as the midpoint of our age range. We next fit a regression model with total vocabulary as our dependent variable, and age as the 1^{st} predictor, number of TEs the 2^{nd} predictor, and the interaction between age and number of TEs the 3^{rd} predictor. We utilized the model-estimated coefficients to compute the weights for TEs at each age. Specifically, we multiplied the estimate for the interaction effect times the age centered variable and added it to the main effect (see the formula below with an example).

Formula for Computing Adjustment Coefficients for TEs

$$\begin{array}{c}\text{Adjustment}\\\text{Coefficients}\end{array} = \begin{array}{c}\text{Effect of}\\\text{\# of TE}\end{array} + \begin{array}{c}\text{Effect of interaction}\\\text{between age and}\\\text{\# of TE}\end{array} \times \text{(Age centered)}$$

Here is an example for a child who is 24 months:

$$1.98 = 1.78 + (-0.06 \text{ x } (24\text{-}25))$$

Following this formula, we computed our adjustment coefficients. Logically, we would expect a range of 1 to 2 for a pair of translation equivalents. Thus, the adjustment coefficients were truncated to 2 and children received a minimum of 1 credit or maximum of 2. This process resulted in the following adjustment coefficients and the truncated adjustment coefficients (see table 2 below). Note these values differ from those reported in Gonzalez-Barrero et al. (2020).

Table 2

Adjustment Coefficients for TEs for Current Study

Age in months	Adjustment Coefficients	Truncated Adjustment Coefficients
18	2.2	2.00
19	2.14	2.00
20	2.08	2.00
21	2.02	2.00
22	1.96	1.96
23	1.9	1.9
24	1.84	1.84
25	1.78	1.78
26	1.72	1.72
27	1.66	1.66
28	1.60	1.60
29	1.54	1.54
30	1.48	1.48
31	1.42	1.42
32	1.36	1.36
33	1.30	1.30

Note. See Table 3 for a comparison of adjustment coefficients in our study to Gonzales et al. (2020).

Using the truncated adjustment coefficients, we calculated the BAV score for each child. The formula to calculate BAV is the sum of the words only known in Spanish, the words known only known in English and the product of the number of TEs times the truncated adjustment coefficients.

Formula for Calculating BAV:

$$\text{BAV} = \begin{array}{c}\text{Words only}\\ \text{known in Spanish}\end{array} + \begin{array}{c}\text{Words only}\\ \text{known in English}\end{array} + \left(\text{\# of TE} \times \begin{array}{c}\text{Truncated}\\ \text{Adjustment}\\ \text{Coefficients}\end{array} \right)$$

Analyses

We conducted analyses following Gonzalez-Barrero et al. (2020) as well. Our analyses focused on how vocabulary size is predicted by age and bilingualism with each of the vocabulary measures. We utilized the lmertest package in R (Kuznetova et al., 2017) and fit linear mixed effects models. This approach allowed us to add random intercepts for participants to account for interdependency due to repeated measures (Bates et al., 2015).

Our models included age as a continuous variable centered at 25 months and language group as a categorical variable with monolinguals as the reference group. Centering age allows us to interpret the model intercept at 25 months which is the midpoint of our age range. Our predictors were age, language group and the interaction between age and language group. Analyses were conducted to address the following question:

1. Does the effect of age, language group or the interaction between age and language group predict vocabulary size across the different vocabulary measures?

Age was included as a predictor as it makes logical sense that older children will have larger vocabularies. This is consistently supported in the literature (Vagh et al., 2009). Language group was an additional predictor because we are interested in seeing if the measures lead to similar vocabulary scores for bilinguals and monolinguals. If so, we do not expect to observe a main effect of language group. The interaction between these two predictors was also of key interest to see if according to any measure, bilinguals have differential vocabulary growth over age.

We expect to replicate past findings in the literature regarding each measure (total, dominant, non-dominant and concept vocabulary). Thus, our analyses allow us to confirm if traditional measures produce similar vocabulary scores across age and/or vocabulary growth rate for bilinguals vs. monolinguals in our sample. In turn, we can test if the BAV measure yields similar vocabulary scores across age and/or vocabulary growth for bilinguals and monolinguals. Furthermore, we can compare the differences captured between monolinguals and bilinguals for each measure in our sample.

Our analyses will allow us to compare the adjustment coefficients and the differences mentioned above for the original study by Gonzalez-Barrero et al. (2020) and for this study. We expect symmetry to play a role in the calculation of the adjustment coefficients. The symmetric sample in Gonzalez-Barrero et al. may promote more opportunities for TE learning compared to our asymmetric sample in which opportunities will be limited by vocabulary in children's nondominant language thus we expect TEs to be a larger part of vocabulary. Given the asymmetric exposure in our sample, children may not have as many opportunities to learn TEs and they may make up less of their vocabulary. If TEs make up a large portion of vocabulary at a certain age, then TEs should count more. This aligns with the expectation of adjustment coefficients to determine how much TEs should count at each age. If so, we predict the symmetric sample to produce comparably larger adjustment coefficients and the asymmetric sample to produce smaller adjustment coefficients.

As a result, we predict the model estimated coefficients will be comparably larger to produce larger adjustment coefficients in Gonzalez-Barrero et al. This means for our dataset; we expect to see smaller estimates and smaller adjustment coefficients. Note larger estimates would be larger, positive values for the main effect and larger, negative values for the interaction. Smaller estimates would be seen as smaller, positive values for the main effect and smaller negative values for the interaction. This is evident in the adjustment coefficients formula. Positive values of the main effect and negative values of the interaction contribute to an increase in the calculated adjustment coefficients. (See Formula 1 for an example). Therefore, we predict the estimates for main effect and interaction will be smaller in our dataset. With a comparison group, we can inspect the findings closely and test the efficacy of this metric.

Results

Traditional Measures

We fit a series of linear mixed-effect models for each of the vocabulary measures with age and language group as a predictor. We discuss the results for the traditional measures before discussing the results with the new metric. For total vocabulary, there was a significant main effect of age. We did not find a main effect of language or interaction with language group. This result is not surprising given the mixed findings with bilinguals appearing to have similar or higher vocabularies when using total vocabulary (Bosch & Ramon-Casas, 2014; Core et al., 2013; Hoff et al., 2013; De Houwer et al., 2014; Gonzalez-Barrero et al., 2020; Junker & Stockman, 2002; Pearson et al., 1993; Thordardottir et al., 2006). All other models with the traditional measures had a significant main effect or interaction with language group.

For dominant and non-dominant vocabulary, there was a statistically significant interaction between age and language group. This outcome paints bilinguals as acquiring vocabulary at a lower rate of growth across age than monolinguals. We see increasingly greater gains in monolingual vocabulary as age increases compared to smaller increases in vocabulary over age in bilinguals. This also aligns with modeling work showing a similar interaction (Bilson et al., 2015). For concept vocabulary, there was a statistically significant main effect of language group. Bilinguals appeared to have lower scores overall than monolinguals. This finding aligns with previous work showing bilinguals having a smaller number of concepts in their vocabulary compared to monolinguals (Core et al., 2013; Gonzalez-Barrero et al., 2020; Thordardottir et al., 2006). In conclusion, most of the traditional measures painted bilingual toddlers' vocabulary in a negative light, in terms of smaller size or slower growth.

BAV

We conducted the same analysis with the BAV metric as the vocabulary measure. With this measure, we observed a significant main effect of age but no main effect or interaction with language group. This is expected as age is a known predictor of vocabulary. This result reflected the ability of this measure to characterize bilingual vocabulary accurately on the assumption that monolinguals and bilinguals should not differ in vocabulary and replicated the findings reported by Gonzalez-Barrero and colleagues (2020). Thus, this approach potentially captures bilingual vocabulary, at least between 18-33 months of age. A limitation of this approach is evident as the mathematical approach does not empirically test whether TEs are harder to learn at earlier ages. We inspected this measure more closely to further evaluate its efficacy (see section Evaluating Efficacy).

Comparison of Adjustment Coefficients

The samples in the original study (Gonzalez-Barrero et al., 2020) and this study are noticeably different. The populations vary largely, particularly in respect to the bilingual populations involved. The original study collected data in Ontario, Quebec, Canada. In this location, English and French are both frequently utilized. French is the official language though English shares an official status. Both languages are taught at school and integrated into the culture. The multicultural status of the city provides both languages a similar status. This setting fits the description of a promotive context, where bilingualism is promoted, and the environment is conducive to bilingual language development (Ortega, 2019). Additionally, the language exposure criterion in this sample ranged from 25% to 75% for both languages over a child's lifetime using the Language Exposure Questionnaire with the MAPLE (Bosch & Sebastian-Galles, 2001; Byers-Heinlein et al., 2008). This sample can be characterized as a symmetric sample, along with consideration of the stringent language exposure criteria used by Gonzalez-Barrero et al.

In contrast, our sample consisted of data collected in Dallas, Texas, United States and available in Wordbank. Arguably, Spanish and English do not hold the same status as the English-French language pair reported in Gonzalez-Barrero et al. Spanish is not recognized as an official language in Texas although it has been integrated into mainstream culture. This setting can be described as a restrictive context, where bilingualism is not promoted, nor is the environment conducive to bilingual language development (Ortega, 2019). These differences can stem from socioeconomic status, percentage of immigrant speakers, and other socio-cultural factors that are harder to assess.

The adjustment coefficients calculated for each study are shown below. These credits reflect the amount of weight a TE should count at each age (derived mathematically). See Table 3. We see evident differences with the values reported in the studies. In Gonzalez-Barrero et al., the range was 2.53-1.03 and the values were truncated from 18-23 months of age. In our study, the range was 2.20-1.30 and the values were truncated from 18-21 months of age. Thus, values vary to a greater degree in Gonzalez-Barrero et al. (2020) than in our dataset which had a narrower range. Accordingly, Gonzalez-Barrero et al. truncated a larger portion of their dataset to 2, meaning 18- to 23-month-old children were given full credit (2) for their TEs. We truncated a smaller portion to 2, so only 18- to 21-month-old children were given full credit (2) for TEs known. This difference could be explained due to expected differences in monolingual French children's vocabulary (about 33 less words compared to age matched English children, as seen in Trudeau et al., 1999). Moreover, French monolingual children's language development is marked by greater syntactic complexity compared to English monolinguals. Thus, making it possible to deduce bilingual children learning French as one of their languages may need TEs to count at the max of 2 for a larger part of their early vocabulary development to account for this difference.

Upon visual inspection, the adjustment coefficients reported in Gonzalez-Barerro et al. show a steeper rate of decrease compared to our coefficients which have a shallower slope (see Figure 1 below). The slope in Gonzalez-Barerro et al. levels out such that by 33 months of age children receive close to 1 credit for TEs. We do not see this pattern in our data. Full credit is given at earlier ages as well, though a slightly shorter range. Yet, at around 33 months of age, about 1.3 of credit should be given.

This suggests sample differences having to do with symmetry are driving the differences of adjustment coefficients between Gonzalez et al. and our study. We predicted symmetry would impact adjustment coefficients via larger estimates. Recall, we expected the symmetric sample to lean towards higher adjustment coefficients to account for the larger possibility of TEs and the asymmetric sample to lean towards slightly lower adjustment coefficients given the lower amount of TEs expected. We found this to be true. Gonzalez et al reported the main effect estimate of 1.83 and an interaction estimate of -0.10. We calculated a main effect estimate of 1.78 and an interaction estimate of -0.06. Thus, as predicted, when these values are plugged into the BAV formula, their adjustment coefficients are larger in Gonzalez-Barrero et al. The differences in the computed adjustment coefficients were expected given prior work citing variability in vocabulary and TEs across language pairs and as a function of language exposure and symmetry. We note that these differences capture the distinct populations in each study: a group of symmetrical bilinguals in Gonzalez-Barrero et al. and a group of asymmetrical bilinguals in this study. This may also be a potential constraint on the applications of this approach. If the BAV measure varies widely between symmetric and asymmetric groups, perhaps that is something to account for in the model to develop more accurate adjustment coefficients to improve its overall accuracy.

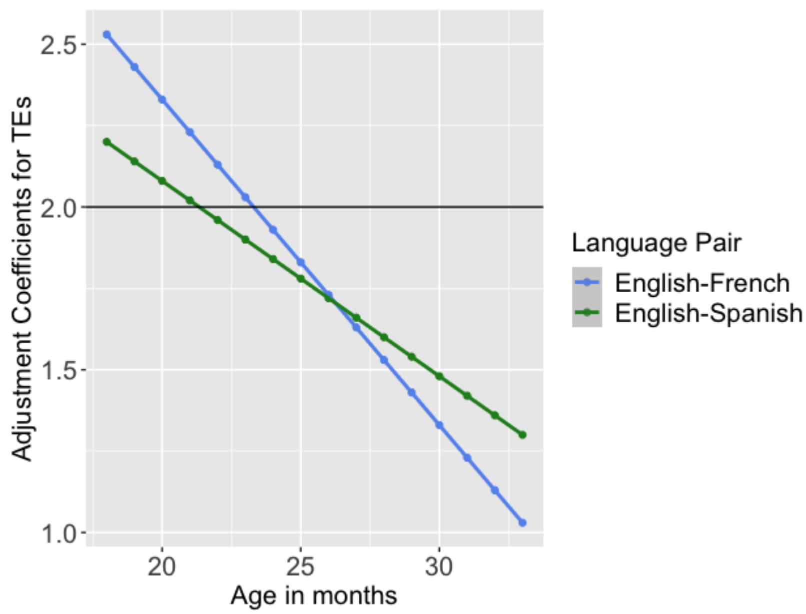
Table 3

Adjustment Coefficients for TEs per Study

Age in months	Gonzalez-Barrero et al. (2020)		Current Study	
	Adjustment Coefficients	Truncated Adjustment Coefficients	Adjustment Coefficients	Truncated Adjustment Coefficients
18	2.53	2.00	2.20	2.00
19	2.43	2.00	2.14	2.00
20	2.33	2.00	2.08	2.00
21	2.23	2.00	2.02	2.00
22	2.13	2.00	1.96	1.96
23	2.03	2.00	1.90	1.90
24	1.93	1.93	1.84	1.84
25	1.83	1.83	1.78	1.78

26	1.73	1.73	1.72	1.72
27	1.63	1.63	1.66	1.66
28	1.53	1.53	1.60	1.60
29	1.43	1.43	1.54	1.54
30	1.33	1.33	1.48	1.48
31	1.23	1.23	1.42	1.42
32	1.13	1.13	1.36	1.36
33	1.03	1.03	1.30	1.30

Figure 1. Adjustment Coefficients for TEs per Language Pair



Evaluating Efficacy of the BAV metric

Given the goal of the BAV metric, it is imperative to evaluate and scrutinize its efficacy. Therefore, we took steps to inspect the metric closely for both the original study (Gonzalez-Barrero et al., 2020) and the current one. First, we inspected the raw means for each vocabulary measure across language groups. We looked closely to see if the BAV measure characterized vocabulary accurately across age groups (see Table 4 below). We split the data into four equal age bins of 3 months each to capture key developments that tend to occur at these ages. Children between the ages of 18-21 months are typically increasing vocabulary at a rapid rate relative to earlier periods of development (Fenson et al., 2007). Children ages 22-25 months begin to produce two-word sentences (Fenson et al., 2007). Children ages 26-29 months will have acquired more vocabulary and are frequently not included in studies with TEs (see David & Wei, 2008; Legacy et al., 2017 for exceptions). The last group of children ages 30-33 captures children who have more word learning experience and have not been always included in vocabulary comparisons (typically up to 30 months; as seen in Core et al., 2013; Pearson et al., 1998). Grouping children by age bands allows us to better compare how BAV characterizes vocabulary in the symmetric sample reported in Gonzalez-Barrero et al. and our asymmetric sample.

As seen in Table 4, the BAV measure overcounts for Gonzalez-Barrero et al. compared to the monolingual vocabulary at 18-21- and 26-29 months and undercounts at ages 30-33 months. In contrast, this measure undercounts compared to monolingual vocabulary in our asymmetric dataset at 18-21 months of age. Crucially, we note why the BAV measure tends to overcount vocabulary in one sample and undercount in the other. The BAV measure stems from evidence derived from the model inspecting the effect of number of TEs and its interaction with age on total vocabulary. If language exposure and age explain the variation in TE learning, the BAV measure should capture these differences in the model-estimated coefficients utilized to compute the adjustment coefficients. BAV overcounts by assigning more credit to TEs in the symmetric sample, which may be more conducive to learning TEs. Thus, giving more credit may not be an appropriate solution. This suggests that accounting for the variation in TE learning does not ultimately correctly quantify bilingual vocabulary in respect to monolinguals.

Additionally, given that the proportion of TEs is expected to increase over age, it would be expected that BAV could overcount for older children, 30-33 months of age. At these ages, bilingual children have acquired a larger vocabulary and may know more TE pairs, at least in comparison to younger bilinguals. This does not align with the current findings. We see the opposite pattern in Gonzalez-Barrero et al. with BAV overcounting at the earliest age group and undercounting in the oldest group. BAV showed subtle numerical differences, either under or overcounting as discussed above. This suggests BAV may need a closer look.

In particular, having 16 different levels of age for each month between the endpoints of the age range tested (18-33 months) makes it difficult to assess how the BAV metric maps onto monolingual vocabulary across more meaningful age ranges. We therefore binned age according to quartiles (18-21, 22-25, 26-29, 30-33) and contrasted monolingual and BAV datasets at each bin using *t*-tests. BAV scores differed from monolingual vocabulary in the Gonzalez-Barrero data set at 18-21 months, $t(161)=25.5778$, $p=0.0012$ and at 30-33 months, $t(112)=2.1135$, $p=0.0368$. Only the 18-21 month bin survived a Bonferroni correction for family-wise error rate, $\alpha=.05/4=.0125$. BAV scores did not differe from monolingual vocabulary in the current data set at any age groups.

If BAV overcounts at statistically significant levels for a symmetric sample but undercounts numerically for an asymmetric one, the combined main effect of number of TEs and its interaction with age does not accurately capture the challenges of learning TEs. Importantly, the presupposed challenge of learning TEs may vary more distinctly across these two samples than this measure can capture.

Table 4

Average Vocabulary Scores per Age Group and Language Group for Gonzalez-Barrero et al. and Current Study (Means and SEs)

	Age Group (months)	Language Group	
		Monolingual	Bilingual
		Actual	BAV
Gonzalez-Barrero et al.	18-21	97 (*10.19*)	146 *(13.33)*
	22-25	319 (*20.09)*	308 *(29.35)*
	26-29	346 *(39.42)*	413 *(28.22)*
	30-33	493 *(14.98)*	442 *(16.29)*
	Age Group (months)	Language Group	
		Monolingual	Bilingual
		Actual	BAV
Current Study	18-21	180 (*15.90)*	138 (*22.20)*
	22-25	236 (*13.60)*	244 (*21.20)*
	26-29	378 *(15.00)*	382 *(43.00)*
	30-33	423 *(26.40)*	425 *(28.80)*

Another criticism of the BAV approach stems from its emphasis on reflecting the inherent difficulty of learning TEs compared to initial words. The BAV metric depends on the assumption that credit for TEs should be assigned so that monolinguals and bilinguals will have the same expected vocabulary size at different ages and thus assigns larger weights (or credits) at younger ages than at older ages. However, the authors do not connect this assumption with an understanding of the cognitive processes inherent in learning initial words vs translation equivalents, processes that may or may not align with projected BAV values. At face value, the method used to calculate adjustment coefficients does not necessarily reflect the difficulty of learning TEs compared to initial words. Although the weights for TEs are higher at younger ages, if these weights reflect difficulty of learning versus the amount of credit we should assign to TEs at younger ages, then perhaps truncating to 2 is not necessary. Importantly, not truncating weights at 2 would amplify the amount of overcounting observed in the Gonzalez-Barrero sample at the youngest age band and align the BAV measure with monolinguals in the current study).

Additionally, Gonzalez-Barrero et al. takes a mathematical approach that tests its predictions on its own data. We followed this approach as well. However, such an approach follows a potentially circular reasoning (Kriegeskorte et al., 2010). Considering the estimates are trained on the data, they are reasonably expected to fare well and accurately predict the expected effect. Yet upon close inspection, the measure is not entirely accurate in predicting vocabulary in bilinguals compared to monolinguals. At the very least. this work should confirm the predictions on additional datasets with characteristics similar to the dataset used for extracting the initial TE weight to fully test its predictions.

In summary, our cross comparison of the BAV predictions for vocabulary across larger age-bands allows us to pinpoint inconsistencies in outcomes using the BAV metric. This measure may over or undercount in comparison to monolingual data. Importantly, it is inconsistent across age groups in an unpredictable manner. The discrepancies across the two samples imply that the BAV approach does not provide a reliable measure. Along with our concerns about the circularity of the approach and the fact that an understanding of the cognitive processes inherent in learning initial words vs translation equivalents may alter assumptions underlying the BAV approach, we highlight the need for empirical study of the differences inherent in learning TEs vs initial words. We describe a potential methodology to address these differences below.

Discussion

In our initial analyses, there was no significant difference between monolinguals and bilinguals when using the BAV measure. Thus, our results provided support for the implementation of BAV to provide comparable norms. However, upon an additional assessment of this data we find subtle differences in how accurately BAV captures vocabulary across different data sets. Given our concern regarding the circularity encompassing this approach, we aim to test the assumptions first and gain valid insight into how this measure can be refined in addition to acquiring more knowledge about how TE learning differs across different bilingual populations.

We question the theoretical assumptions underlying this approach and highlight multiple assumptions that remain unresolved. The first assumption is that bilingual and monolingual children rely on the same word learning mechanisms. This does not quantify the additional differences in the environment that may impact the cognitive processes children use in acquiring language. As noted earlier, the specific context, e.g., the language mode of the experimental task and other context cues, can impact the way these mechanisms are employed (Kalashnikova et al., 2018; Weatherhead et al., 2021). If bilingual children must employ these mechanisms differently, we might expect bilingual children to have different vocabulary outcomes than monolingual children as opposed to similar ones. Interestingly, monolingual children also differ on multiple facets of language because of the specific language they are learning (Mani & Plunkett, 2007). Sensitivity to prosodic cues can differ as seen by the trochaic bias emerging around 6 months of age in English and German learning infants but not those learning French (Hoehle et al., 2009).

Differences in languages with complex sound structure (Danish) or syntactic complexity (French) are assumed to predict vocabulary differences seen with Danish monolinguals showing lower receptive vocabulary comparably to other monolinguals (Bleses et al., 2008) and French monolinguals having smaller vocabularies compared to age-matched English monolinguals despite showing greater syntactic complexity (Thordardottir, 2005). Monolingual children may be relying on language-specific abilities that support their language development. Thus, the assumption that bilingual vocabulary must map onto monolingual levels may be unsubstantiated. Moreover, this highlights the need to include additional factors into this measure to differentiate between subsets of populations given the differential impact of certain factors.

The second assumption we discuss is based on distinct accounts for TE learning. In this section, I refer to TEs as the second label learners acquire and initial words as the first label acquired for a referent (Umbel et al., 1992). Table 5 lists the current accounts of learning of TEs: the preference, neutral and avoidance accounts (Tsui et al., 2022). The preference account refers to an inclination towards learning TEs. Specifically, this account posits a child is more likely to learn a TE, the second label for a referent, rather than an initial word for a referent or concept, for which the child has not yet acquired any labels. The neutral account refers to a lack of inclination towards TEs with no difference in learning a TE or initial word. This account suggests that a child is equally likely to learn a TE, the second label for a referent, as the initial word for a referent for which the child has not acquired any labels. The avoidance account refers to an aversion towards learning TEs. Under this account, a child is less likely to learn a TE, the second label for a referent, compared to an initial, or new label for a referent, for which the child has not yet acquired any labels.

Table 5

Descriptions of TE Accounts

Account	Preference	Neutral	Avoidance
Description	When a TE is easier to learn compared to an initial word	When a TE is equally likely to be learned compared to an initial word	When a TE is harder to learn compared to an initial word

Importantly, these accounts suggest differential patterns of vocabulary development. The preference account implies a larger majority of a child's vocabulary is made up of TEs given they are easier to learn. The neutral account assumes vocabularies are developed independently so the number of TEs and initial words does not pattern predictably. The avoidance account surmises TEs are a minimal part of children's vocabulary if at all, given they are harder to learn. These accounts are the main ones describing how TEs are acquired.

Past research provides some evidence to support or negate each account. We summarize some of the main findings in the literature (see Table 6 below). At a glance, most evidence supports the preference or neutral accounts. Evidence for the preference account includes research supporting shared semantic representations between languages that would mean that acquiring a TE would require only mapping a label onto a referent (Floccia et al., 2020; Francis, 2020; Jardak & Byers-Heinlein, 2019; Singh, 2014). This would facilitate the acquisition of the TE given the existing semantic representation. Similarly, Bilson and colleagues demonstrated via a computational model that bilingual children's acquisition of vocabulary follows a predictable pattern where they are more likely to acquire words that relate semantically to known words in their lexicons supporting a preference account. Also following a computational modeling approach to predict various scenarios for TEs according to each account, Tsui and colleagues found support for the preference account at smaller vocabulary sizes and the neutral account at larger vocabulary sizes (2022). In addition, bilinguals learn more TEs when their dominant and non-dominant language switch (Lanvers, 1999; Pearson & Fernandez, 1994) presumably stemming from increased exposure to the former non-dominant language which should increase opportunities for learning new TEs. Perhaps, this moment of language switch could be characterized as an overall push towards symmetry.

The neutral account does not lean in favor of learning TEs or initial words. As noted, Tsui et al. (2022) found evidence for the neutral account at larger vocabulary sizes. Random matching between the separate vocabularies of Spanish and English monolingual children to showed similar overlaps in TEs compared to those of actual bilingual children (Pearson et al., 1995), hinting that there is possibly no preference for TEs. Furthermore, processing efficiency in bilingual toddlers correlates to vocabulary size in the same language, but not to the alternate language's vocabulary size (Marchman et al., 2010). Furthermore, large differences in vocabulary growth across languages of bilingual children are common depending on relative exposure to their two languages (Pearson & Fernandez, 1994; Pearson et al., 1997). Thus, the lack of association between bilinguals' two languages suggests their development is independent of one another and would not contribute to learning TEs or initial words over another.

Most evidence for the avoidance account comes from early theories of bilingual language development (Imedadze, 1967; Swain & Wesche, 1975; Volterra & Taeschner, 1978). This includes the assumption that bilingual children are unable to differentiate between their languages. However, bilingual children adapt or switch the language they utilize according to the language employed in the current environment (Paradis and Nicoladis, 2007) or with respect to a particular speaker (Genesee et al., 1995, Genesee et al., 1996). Though their languages interact, bilingual children can differentiate their languages (Byers-Heinlein & Lew-Williams, 2013). Similarly, a word-learning heuristic, mutual exclusivity, was initially cited as support for this idea (Markman, 1994). In principle, mutual exclusivity fits with this account. This word learning strategy suggests that children will assume each object has one label which would imply that bilingual children would not be inclined to adopt an additional label (TE) for an object for which they already have one label. However, bilingual children violate this principle which is unreliable given their language learning experiences (Byers-Heinlein et al., 2013. Houston-Price et al., 2010; Weatherhead et al., 2021). Thus, the ideas that support this account lack strong, current evidence.

Summing across these studies, none directly address and compare learning of TEs vs initial words which would provide direct evidence for resolving these accounts. Indeed, the findings summarized in Table 6 provide only indirect evidence for these accounts that remain unresolved.

Table 6

Summary of empirical support for existing accounts of TE

Evidence	Relevant Concepts & Findings	Preference	Neutral	Avoidance
Bilson et al., 2015	Semantically related words to bilingual children's lexicon were learned faster	✓		
Floccia et al., 2020; Francis, 2020; Jardak & Byers-Heinlein, 2019; Singh, 2014	Shared semantic representations between languages	✓		

Imedadze, 1967; Swain & Wesche, 1975; Volterra & Taeschner, 1978	Early theories of bilingual language development: lack of language differentiation			✓
Markman, 1994	Mutual exclusivity principle: one label to one object mapping			✓
Lanvers, 1999; Pearson & Fernandez, 1994	Bilingual children learned TEs more when their language exposure shifted their dominant and non-dominant languages	✓		
Marchman et al., 2010	Bilingual children's processing efficiency in one language is correlated to vocabulary size in that language but not in the other language		✓	
Pearson & Fernandez, 1994; Pearson et al., 1997	Unequal vocabulary growth in both languages due to language exposure differences		✓	
Pearson et al., 1995	TEs were found to be equally acquired		✓	
Tsui et al., 2022	Computational modeling: a shift from preference account to neutral	✓	✓	

As discussed above, existing evidence supports two of the accounts, but the evidence does not relate directly to how learning of TEs compare to initial words and if that varies across age. Additionally, only Tsui et al. points to a shift between accounts or when during development a shift may occur. However, the evidence presented in this paper uses the same dataset in Gonzalez-Barrero et al., (2020) to confirm their simulated data predictions for each TE account, limiting its implications. Given the evidence available, the preference account may not be an accurate representation for all bilingual children, depending on various factors such as languages involved, language exposure, and context of language learning.

Hence, we argue for future research to provide concrete evidence for the preference account or other TE accounts. Doing so will allow us to fill the gaps in the theoretical accounts of vocabulary development and help refine appropriate measures. The existing evidence points to a need to address the assumptions and additional inclusions in this measure to refine it. This may allow us to adjust the key factors we may wish to consider in refining approaches to predicting vocabulary in different kinds of bilinguals. Language exposure and language dominance is not enough to consider. Perhaps these variables need to be teased apart into language symmetry and the degree of asymmetry. It may be a greater cognitive challenge for asymmetric bilinguals to learn TEs. This may show up as difference in the number of TEs known even with some exposure to the same objects in both languages. Moreover, this may explain the differences seen in bilingual children appearing to know words in each vocabulary depending on the context (Hoff et al., 2014). Hoff and colleagues posited this context-dependent development was predicted by language exposure. Yet more recent work has shown vocabulary development may be impacted by the child’s usage of the languages (Hoff, 2020). Accounting for language symmetry and degree of asymmetry may be a way to capture or explain exposure and usage. This is one way research can expand upon this measure.

In conclusion, our results after applying the BAV measure to symmetrical vs. asymmetrical populations lead us to question the theoretical and empirical foundations of the BAV measure. We must have a stronger understanding of the different challenges children face in learning TEs vs. IWs as part of their vocabulary development. Specifically, we should aim to understand if and how those differences vary in symmetric vs asymmetric bilingual populations. Until we establish better understanding, we are limited in developing a new measure that better accounts for bilingual vocabulary, in terms of comparable norms. More work is required before we can garner insight into how bilingual children learn vocabulary and more broadly, into how they acquire their languages.

Appendix A

Adjustment Coefficients for TEs per Study

Age in months	Gonzalez-Barrero et al. (2020)		Current Study	
	Adjustment Coefficients	Truncated Adjustment Coefficients	Adjustment Coefficients	Truncated Adjustment Coefficients
18	2.53	2.00	2.20	2.00
19	2.43	2.00	2.14	2.00
20	2.33	2.00	2.08	2.00
21	2.23	2.00	2.02	2.00
22	2.13	2.00	1.96	1.96
23	2.03	2.00	1.90	1.90
24	1.93	1.93	1.84	1.84
25	1.83	1.83	1.78	1.78
26	1,73	1,73	1.72	1.72
27	1.63	1.63	1.66	1.66
28	1.53	1.53	1.60	1.60
29	1.43	1.43	1.54	1.54
30	1.33	1.33	1.48	1.48
31	1.23	1.23	1.42	1.42
32	1.13	1.13	1.36	1.36
33	1.03	1.03	1.30	1.30

References

Alt, M., Arizmendi, G. D., Gray, S., Hogan, T. P., Green, S., & Cowan, N. (2019). Novel word learning in children who are bilingual: Comparison to monolingual peers. *Journal of Speech, Language, and Hearing Research*, *62*(7), 2332–2360. https://doi.org/10.1044/2019_jslhr-l-18-0009

Alt, M., Gray, S., Hogan, T. P., Schlesinger, N., & Cowan, N. (2019). Spoken word learning differences among children with dyslexia, concomitant dyslexia and developmental language disorder, and typical development. *Language, Speech, and Hearing Services in Schools*, *50*(4), 540–561. https://doi.org/10.1044/2019_lshss-voia-18-0138

Arizmendi, G. & Alt, M. (in prep). The development and validation of a novel task to quantify functional language proficiency in Spanish English learning school-age children.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using**lme4**. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bedore, L. M., & Peña, E. D. (2008). Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism*, *11*(1), 1–29. https://doi.org/10.2167/beb392.0

Bialystok, E., Luk, G., Peets, K.F, & Yang, S.U. (2009). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *13*(4), 525–531. https://doi.org/10.1017/s1366728909990423

Bijeljac-Babic, R., Höhle, B., & Nazzi, T. (2016). Early prosodic acquisition in bilingual infants: The case of the perceptual trochaic bias. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00210

Bleses, D., Højen, A., Justice, L. M., Dale, P. S., Dybdal, L., Piasta, S. B., Markussen-Brown, J., Clausen, M., & Haghish, E. F. (2017). The effectiveness of a large-scale language and preliteracy intervention: The spell randomized controlled trial in Denmark. *Child Development*, *89*(4). https://doi.org/10.1111/cdev.12859

Borovsky, A., Ellis, E.M., Evans. J.L., & Elman, J.L. (2016). Lexical leverage: category knowledge boosts real-time word recognition in 2-year-olds. *Developmental Science, 19*(6), 918-932. doi: 10.1111/desc.12343.

Bosch, L., & Ramon-Casas, M. (2014). First translation equivalents in bilingual toddlers' expressive vocabulary. *International Journal of Behavioral Development*, *38*(4), 317–322. https://doi.org/10.1177/0165025414532559

Byers-Heinlein, K., Fennell, C.T., & Werker, J.F. (2013). The development of associative word learning in monolingual and bilingual infants. *Bilingualism: Language and Cognition, 16*, 198-205. doi: 10.1017/S1366728912000417

Byers-Heinlein, K., & Lew-Williams, C. (2013). Bilingualism in the Early Years: What the Science says. *LEARNing Landscapes*, *7*(1), 95–112. https://doi.org/10.36510/learnland.v7i1.632

Carter, A., & Gerken, L.A (1997). Functors in Early On-Line Sentence Comprehension. *The Proceedings of The Twenty-Eight Annual Child Language Research Forum*. 53-61.

Cattani, A., Abbot-Smith, K., Farag, R., Krott, A., Arreckx, F., Dennis, I., & Floccia, C. (2014). How much exposure to English is necessary for a bilingual toddler to perform like a monolingual peer in language tests? *International Journal of Language & Communication Disorders*, *49*(6), 649–671. https://doi.org/10.1111/1460-6984.12082

Core, C., Hoff, E., Rumiche, R., & Señor, M. (2013). Total and conceptual vocabulary in Spanish–English bilinguals from 22 to 30 months: Implications for assessment. *Journal of Speech, Language, and Hearing Research*, *56*(5), 1637–1649. https://doi.org/10.1044/1092-4388(2013/11-0044

Crutchley, A., (2000). Bilingual children in language units: does having ‘well-informed parents’ make a difference? International Journal of Language and Communication Disorders, 35, 65–81.

Dale, P. S. (1991). The validity of a parent report measure on vocabulary and syntax at 24 months. *Journal of Speech & Hearing Research, 34*(3), 565–571. https://doi.org/10.1044/jshr.3403.565

David, A., & Wei, L. (2008). Individual differences in the lexical development of French–English bilingual children. *International Journal of Bilingual Education and Bilingualism*, *11*(5), 598–618. https://doi.org/10.1080/13670050802149200

De Houwer, A., Bornstein, M. H., & Putnick, D. L. (2014). A bilingual–monolingual comparison of young children's vocabulary size: Evidence from comprehension and production. *Applied Psycholinguistics*, *35*(6), 1189–1211. doi:10.1017/S0142716412000744

Duff, F. J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry*, *56*(8), 848–856. https://doi.org/10.1111/jcpp.12378

Ellis, E. M., & Thal, D. J. (2008). Early language delay and risk for language impairment. *Perspectives on Language Learning and Education*, *15*(3), 93–100. https://doi.org/10.1044/lle15.3.93

Fenson, L., Marchman, V. A., Thal, D., Dale, P. S., & Bates, E. (2007). *MacArthur - Bates Communicative Development Inventories (CDIs)* (2nd ed.). Baltimore: Brookes Publishing.

Francis, W. S. (2020). Shared core meanings and shared associations in bilingual semantic memory: Evidence from research on implicit memory. *International Journal of Bilingualism, 24*(3), 464-477. doi:10.1177/1367006918814375

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2016). Wordbank: An open repository for developmental vocabulary data. Journal of Child Language. https://doi.org/1017/S0305000916000209

Genesee, Fred & Nicoladis, Elena & Paradis, Johanne. (1995). Language differentiation in early bilingual development. Journal of child language. 22. 611-31. doi:10.1017/S0305000900009971.

Genesee, Fred & Boivin, Isabelle & Nicoladis, Elena. (1996). Talking with strangers: A study of bilingual children's communicative competence. Applied Psycholinguistics. 17. 427 - 442. doi:10.1017/S0142716400008183.

Gerken, L. A., & McIntosh, B. J. (1993). Interplay of func- tion morphemes and prosody in early language. Developmental Psychology, 29, 448-457.

Gonzalez-Barrero, A.M., Schott, E., & Byers-Heinlein, K. Bilingual adjusted vocabulary: A developmentally-informed bilingual vocabulary measure. PsyArXiv, 2020. 7, 2020.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language, 36*(1), 3–15. https://doi.org/10.1016/0093-934X(89)900485

Hoff, E. (2010). Context effects on young children's language use: The influence of conversational setting and partner. *First Language*, *30*(3-4), 461–472. https://doi.org/10.1177/0142723710370525

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. Journal of Child Language, 39(01), 1-27.

Hoff, E., Rumiche, R., Burridge, A., Ribot, K. M., & Welsh, S. N. (2014). Expressive vocabulary development in children from bilingual and monolingual homes: A longitudinal study from two to four years. *Early Childhood Research Quarterly*, *29*(4), 433–444. https://doi.org/10.1016/j.ecresq.2014.04.012

Houston-Price, C., Caloghiris, Z., & Raviglione, E. (2010), Language experience shapes the development of the mutual exclusivity bias. *Infancy, 15* (2), 125-150. https://doi.org/10.1111/j.1532-7078.2009.00009.x

Jackson-Maldonado, D., Thal, D., Marchman, V., Bates, E., & Gutierrez-Clellen, V. (1993). Early lexical development in Spanish-speaking infants and toddlers. *Journal of Child Language, 20*(3), 523–549. https://doi.org/10.1017/S0305000900008461

Junker, D. A., & Stockman, I. J. (2002). Expressive vocabulary of German-English bilingual toddlers. *American Journal of Speech-Language Pathology*, *11*(4), 381–394. https://doi.org/10.1044/1058-0360(2002/042)

Kachergis, G., Marchman, V. A., Dale, P., Mankewitz, J., & Frank, M. C. (2021, July 6). Online Computerized Adaptive Tests (CAT) of Children's Vocabulary Development in English and Mexican Spanish. https://doi.org/10.31234/osf.io/5ftsu

Kalashnikova, M., Escudero, P., & Kidd, E. (2018). The development of fast-mapping and novel word retention strategies in monolingual and bilingual infants. *Developmental Science*, *21*(6). https://doi.org/10.1111/desc.12674

Kriegeskorte, N., Lindquist, M. A., Nichols, T. E., Poldrack, R. A., & Vul, E. (2010). Everything you never wanted to know about circular analysis, but were afraid to ask. *Journal of Cerebral Blood Flow & Metabolism*, *30*(9), 1551–1557. https://doi.org/10.1038/jcbfm.2010.86

Kohnert, K. (2010). Bilingual children with primary language impairment: Issues, evidence and implications for clinical actions. Journal of Communication Disorders, 43(6), 456-473.

Lanvers, U. (1999). Lexical growth patterns in a bilingual infant: The occurrence and significance of equivalents in the bilingual lexicon. *International Journal of Bilingual Education and Bilingualism*, *2*(1), 30–52. https://doi.org/10.1080/13670059908666245

Law, J., & Roy, P. (2008). Parental report of infant language skills: A review of the development and application of the communicative development inventories. *Child and Adolescent Mental Health*, *13*(4), 198–206. https://doi.org/10.1111/j.1475-3588.2008.00503.x

Legacy, J., Reider, J., Crivello, C., Kuzyk, O., Friend, M., Zesiger, P., & Poulin-Dubois, D. (2017). Dog or chien? Translation equivalents in the receptive and expressive vocabularies of young French–English bilinguals. *Journal of Child Language*, *44*(04), 881–904. https://doi.org/10.1017/S0305000916000295

Mancilla-Martinez, J., & Lesaux, N. K. (2011). The gap between Spanish speakers' word reading and word knowledge: A longitudinal study. *Child Development*, *82*(5), 1544–1560. https://doi.org/10.1111/j.1467-8624.2011.01633.x

Mani, N., & Plunkett, K. (2010). In the infant's Mind's ear. *Psychological Science*, *21*(7), 908–913. https://doi.org/10.1177/0956797610373371

Marchman, V. A. (1999). Scoring program for the MacArthur–Bates Communicative Development Inventories. Retrieved from http://www.sci.sdsu.edu/cdi/scoringdbp.html

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*(3). https://doi.org/10.1111/j.1467-7687.2008.00671.x

Marchman, V. A., & Martínez-Sussman, C. (2002). Concurrent validity of caregiver/parent report measures of language for children who are learning both English and

Spanish. *Journal of Speech, Language, and Hearing Research, 45*(5), 983–997. https://doi.org/10.1044/1092-4388(2002/080)

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*(2), 212–224. https://doi.org/10.1111/j.1467-7687.2004.00340.x

Nicoladis, E., & Genesee, F. (1997). Language development in preschool bilingual children. *Journal of Speech-Language Pathology and Audiology, 21*(4), 258–270.

Nicoladis, E., & Wiebe, S. A. (2019). How to use a wide variety of words in telling a story with a small vocabulary: Cognitive predictors of lexical selection for simultaneous bilingual children. *Language, Cognition and Neuroscience*, *35*(3), 330–338. https://doi.org/10.1080/23273798.2019.1678758

O'Toole, C., Gatt, D., Hickey, T. M., Miękisz, A., Haman, E., Armon-Lotem, S., Rinker, T., Ohana, O., dos Santos, C., & Kern, S. (2016). Parent report of early lexical production in bilingual children: A cross-linguistic CDI comparison. *International Journal of Bilingual Education and Bilingualism*, *20*(2), 124–145. https://doi.org/10.1080/13670050.2016.1179258

Ortega, L. (2019). The study of Heritage Language Development from a bilingualism and Social Justice Perspective. *Language Learning*, *70*(S1), 15–53. https://doi.org/10.1111/lang.12347

Paradis, J., Crago, M., Genesee, F., & Rice, M. (2003). French-English bilingual children with SLI: How do they compare with their monolingual peers? Journal of Speech, Language, and Hearing Research, 46(1), 113-127.

Paradis, J. & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or interdependent? Studies in Second Language Acquisition, 18, 1-25.

Pearson, B. Z., Fernandez, S. C., Lewedeg, V., & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, *18*(1), 41–58. https://doi.org/10.1017/s0142716400009863

Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning*, *43*(1), 93–120. https://doi.org/10.1111/j.1467-1770.1993.tb00174.x

Poulin-Dubois, D., Bialystok, E., Blaye, A., Polonia, A., & Yott, J. (2013). Lexical access and vocabulary development in very young bilinguals. *International Journal of Bilingualism*, *17*(1), 57–70. https://doi.org/10.1177/1367006911431198

Rocha-Hidalgo, J., Feller, M., Blanchfield, O. A., Kucker, S. C., & Barr, R. F. (2021). Patterns of mutual exclusivity and Retention: A Study of Monolingual and bilingual 2-year-olds. *Infancy*, *26*(6), 1011–1036. https://doi.org/10.1111/infa.12432

Schonberg, C. C., Russell, E. E., & Luna, M. L. (2019). Effects of past language experience and present language context on the shape bias in Spanish–English bilingual children. *Developmental Science*, *23*(2). https://doi.org/10.1111/desc.12879

Setoh, P., Cheng, M., Bornstein, M. H., & Esposito, G. (2021). Contrasting lexical biases in bilingual English–Mandarin speech: Verb-biased mothers, but noun-biased toddlers.

Journal of Child Language, *48*(6), 1185–1208. https://doi.org/10.1017/s0305000920000720

Storkel, H. L., Bontempo, D. E., Aschenbrenner, A. J., Maekawa, J., & Lee, S. Y. (2013). The effect of incremental changes in phonotactic probability and neighborhood density on word learning by preschool children.

Thordardottir E. Early lexical and syntactic development in Quebec French and English: Implications for cross-linguistic and bilingual assessment. International Journal of Language and Communication Disorders 2005; 40: 243–278

Thordardottir, E., Rothenberg, A., Rivard, M.-E., & Naves, R. (2006). Bilingual assessment: Can overall proficiency be estimated from separate measurement of two languages? *Journal of Multilingual Communication Disorders*, *4*(1), 1–21. https://doi.org/10.1080/14769670500215647

Trudeau, N., Frank, I., & Poulin-Dubois, D. (1999). Une adaptation en français québecois du MacArthur Communicative Development Inventory [A Quebec French adaptation of the MacArthur Communicative Development Inventory]. *Revue d'orthophonie et d'audiologie*, *23*, 31–73.

Tsui, R. K., Gonzalez-Barrero, A. M., Schott, E., & Byers-Heinlein, K. (2022). Are translation equivalents special? evidence from simulations and empirical data from bilingual infants. *Cognition.* https://doi.org/10.31234/osf.io/h4jxe

Uccelli, P., & Páez Mariela 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*(3), 225–236. https://doi.org/10.1044/0161-1461(2007/024)

United States Census Bureau. (2017). *Selected social characteristics in the United States, 2017 American community survey 5-year estimates.* Retrieved from https://data.census.gov/cedsci/table?d=ACS%205-Year%20Estimates%20Data%20Profiles&tid=ACSDP5Y2017.DP02.

Vagh, S. B., Pan, B. A., & Mancilla-Martinez, J. (2009). Measuring growth in bilingual and monolingual children's English productive vocabulary development: The utility of combining parent and teacher report. *Child Development*, *80*(5), 1545–1563. https://doi.org/10.1111/j.1467-8624.2009.01350.x

Weatherhead, D., Kandhadai, P., Hall, D. G., & Werker, J. F. (2021). Putting mutual exclusivity in context: Speaker race influences monolingual and bilingual infants' word-learning assumptions. *Child Development*, *92*(5), 1735–1751. https://doi.org/10.1111/cdev.13626

Werker, J. F., & Byers-Heinlein, K. (2008). Bilingualism in infancy: First steps in perception and comprehension. *Trends in Cognitive Sciences*, *12*(4), 144–151. https://doi.org/10.1016/j.tics.2008.01.008