

METHODS FOR TARGETING VOCABULARY DEVELOPMENT IN
SECOND LANGUAGE LEARNERS:
A SYSTEMATIC REVIEW

by

Linnea E. Larson

Submitted in partial fulfillment of the
requirements for Departmental Honors in
the Department of Communication Sciences and Disorders

Texas Christian University

Fort Worth, Texas

May 2, 2014

METHODS FOR TARGETING VOCABULARY DEVELOPMENT IN
SECOND LANGUAGE LEARNERS:
A SYSTEMATIC REVIEW

Project Approved:

Supervising Professor: Christopher Watts, Ph.D.

Department of Communication Sciences and Disorders

Lynita Yarbrough, M.S. - EBSLP

Department of Communication Sciences and Disorders

Danielle Brimo, Ph.D.

Department of Communication Science and Disorders

ABSTRACT

The purpose of this review is to **(a)** identify which interventions have been used over the last decade for targeting vocabulary development in second language learners where the first language is typically Spanish, **(b)** to evaluate the methodological rigor in which those interventions were investigated, and **(c)** determine the clinical relevance of the outcome measurements used to judge if the intervention(s) were efficacious. This was achieved by performing a systematic review both identifying and describing techniques for enhancing vocabulary growth in second language learners. Five studies were included in the systematic review and fell within a ten-year range. It was found that the most effective strategies used a combination of techniques and relied heavily on reading which the least effective did not incorporate any sort of reading strategy.

TABLE OF CONTENTS

| | |
|--|----|
| INTRODUCTION | 1 |
| Theories of Vocabulary Learning | 2 |
| PURPOSE | 7 |
| METHODS | 8 |
| Search Strategies | 8 |
| Filtering of Studies | 9 |
| Table 1: Levels of Evidence Used for Inclusion Criteria | 10 |
| RESULTS | 11 |
| Interventions for Targeting Vocabulary Development | 11 |
| Table 2a: Characteristics of Interventions for Included Studies | 13 |
| Table 2b: Dependent Variables Investigated in Each Included Study | 14 |
| Table 2c: Experimental Design and Participant Characteristics of Studies | 15 |
| First Language First Model (Schwartz, 2013) | 16 |
| Educator-Implemented Intervention (Vuattoux et al., 2013) | 17 |
| Conversation Intervention (Ruston & Schwanenflugel, 2010) | 19 |
| Rich Explanation through Storybook Reading (Collins, 2010) | 20 |
| Read-Aloud (Neugebauer & Currie-Rubin, 2009) | 22 |
| Quality of Study Designs | 24 |
| Efficacy of Intervention | 24 |
| Table 3: Calculated Effect Sizes for Experimental and Control Groups | 25 |
| Discussion | 26 |
| SUMMARY | 28 |

REFERENCES30

INTRODUCTION

Early exposure to books and participation in joint book reading contribute to increased phonemic awareness, print awareness, readiness for school, and educational gains as children enter learning communities (Brooks, 2010; Bus, Ljzendoorn van & Pellegrini, 1995; Justice & Ezell, 2002; Owens, 2010). However, despite the clear advantages of reading and widespread availability of books within the United States, 2.3% of elementary school children experience academic failure due to underdeveloped reading skills (West, 2012). One primary cause of this reading-related failure in the early years of school can be attributed to a lack of vocabulary breadth and depth (Biemiller, 2006). Breadth of vocabulary is defined as the number of words a person knows that hold at least a superficial meaning—or put simply, the size of a person’s vocabulary, or their lexicon. Depth of vocabulary refers to the quality of vocabulary knowledge or the “level of knowledge of various aspects of a given word” (Shen, 2008). These two components together make up the overarching term of vocabulary (semantics).

According to the National Center for Education Statistics (2012), Hispanic and African American children’s vocabulary scores were lower when compared to White and Asian/Pacific Islander (i.e., almost 4 times the rate of White children). Furthermore, norm-referenced vocabulary assessments have shown that bilingual children tend to perform at a lower level than monolingual peers and often times at a level comparable to “monolingual language-impaired children” (Dufrense & Masny, 2006). Evidence from researchers has suggested that low socioeconomic status, below average phonemic awareness ability, extreme stress, under-stimulation, and little positive language

contribute to children's' reduced vocabulary breadth and depth (Lady of the Lake University, 2012; Hart & Risely, 1995).

Language acquisition theories have been developed that support the notion that reading may be supportive for gains in vocabulary breadth and depth. One such hypothesis, the Input Hypothesis originally developed by Stephen Krashen et al., (1989), assumes that we acquire language through comprehension of messages, or a listener's understanding of the message. Krashen argued that competence in vocabulary is most efficiently developed through reading. Competence in vocabulary was defined as the ability to effectively and efficiently use vocabulary. Previous studies have supported this theory by providing evidence for a high mutually-interdependent correlation between reading and vocabulary (Chiappe, Chiappe, & Gottardo, 2004; Biemiller, 2006; Eldredge, J.L., Quinn, B., & Butterfield, D.D., 1990; Oulette, 2006). Therefore, the purpose of this is to investigate the effects of a reading intervention on a bilingual child's vocabulary breadth and depth.

Theories of Vocabulary Learning

Though evidence supports the notion that both depth and breadth of vocabulary knowledge significantly impact reading comprehension, there remains ambiguity in regards to the "best way" to acquire this vocabulary knowledge. In Stephen Krashen's study of the Input Hypothesis, he argues that the best way to learn is not only through "comprehensible input" but through comprehensible input in the form of reading. There are two other hypotheses widely accepted (perhaps more widely accepted in the traditional learning model). The first of these is called the Skill-Building Hypothesis (Krashen, 1989). It suggests that one learns language through first learning individual

rules or items and then making the rules “automatic” through drills and exercises. This theory suggests a “top-down” approach in that vocabulary knowledge starts specific to an item and then generalizes across the individual learned skills. The second hypothesis is the Output Hypothesis. According to this hypothesis, speech production is the catalyst for language learning. It claims that the rules governing language are discovered through attempts at production. Thus, the Output Hypothesis differs from the Skill-Building Hypothesis in that it represents a “bottom-up” method (or inductive learning) (Krashen, 1989).

Studies investigating the Input Hypothesis through reading have shown this theory stimulates vocabulary breadth and depth more efficiently than either the Skill-Building Hypothesis or Output Hypothesis (Miller, 1941; Krashen, 1989). In fact, in one instance, junior high students were asked to participate in a month’s study of natural resources and were not taught any vocabulary prior to or during the process. By the end of one month, Miller remarked that after “the extensive reading done on the subject, our pupils had incidentally accumulated an unusual store of conversation terms...” Thus, her conclusion was that “extensive reading by pupils having definite information goals ahead is most conducive to vocabulary growth”. In other words, though the students were not being taught specific vocabulary words, prefixes, or definitions and were not merely trying to produce certain vocabulary words to better understand them, they expanded both their vocabulary breadth and depth through research and motivation by special interest (Miller, 1941).

The input hypothesis puts forth the idea that vocabulary acquisition is impacted significantly when a child is exposed to more advanced language input than their current

level (comprehensible input). Krashen argues that the best way to facilitate this input is through reading. Literature described above has supported this supposition and confirmed that reading does in fact influence vocabulary breadth and depth positively. However, the input hypothesis—specifically in regards to reading intervention—has not yet been investigated using a specific set of predetermined words. Rather, vocabulary has been assessed and directly targeted through use of simultaneous written and oral language or through “information goals”. In addition to a lack of specific vocabulary targets, there is also very little—if any—research to support the hypothesis in bilingual (Spanish-English speaking) children despite the fact that they may be most at risk for vocabulary delay. Thus, it remains significant to investigate the pattern of vocabulary development secondary to receiving a reading-based vocabulary intervention in a bilingual Spanish-English child who has been identified as delayed in vocabulary knowledge.

Vocabulary and Reading-related skills

In a research article by Gene P. Ouellette (2006), links between written and oral language were explored with specific regards to breadth and depth of vocabulary and reading related skills. Specifically, Ouellette aimed to investigate the link between breadth and depth of vocabulary (i.e., oral language) and reading fluency and reading comprehension abilities (i.e., written language). The study was conducted using a sample of 85 grade four students—60 of whom were used for statistical purposes. Students in the fourth grade were targeted because they were deemed to possess efficient word-level reading skills. Ouellette defined vocabulary growth as “adding and refining phonological representations to the lexicon as well as storing and elaborating the associated semantic knowledge” (Ouellette, 2006). In this study, breadth of vocabulary was measured by

testing receptive and expressive vocabulary through picture identification. Depth of vocabulary was measured by requiring children to select synonyms of a target word and provide a definition for a target word. Word-level recognition and decoding were assessed with an experimental word list adapted from Adams and Huggins (1985) and the Word Attack subtest of the Woodcock Reading Mastery Tests—Revised (Woodcock, 1998), respectively. For word-level recognition, students were required to read aloud from the word list aloud. For word-level decoding, students were required to read pseudowords and were scored based on regular phonics words. Reading comprehension was assessed through the Passage Comprehension subtest of the Woodcock Reading Mastery Tests—Revised. (Woodcock, 1998) Ouellette hypothesized that breadth and depth of oral vocabulary is linked to decoding ability of written words, and that the depth of the vocabulary is crucial to support adequate reading comprehension ability.

The results of this study revealed a significant relation between vocabulary breadth and reading related skills. Word-level decoding was significantly correlated with receptive vocabulary breadth (i.e., $r = .303$). Significant correlations were not found between word-level decoding and expressive vocabulary breadth or depth. Word-level recognition was significantly correlated with receptive and expressive vocabulary breadth and vocabulary depth (i.e., $r = .439, .365, .298$, respectively). Reading comprehension was also significantly correlated with breadth and depth of vocabulary (i.e., $r = .484, .362, .504$, respectively). The results support Ouellette's hypothesis which suggested an association between oral vocabulary and word-level reading and reading comprehension skills. The results of this study support the relation between oral language and written language. Based on the results, Ouellette suggested that to facilitate better

reading comprehension, the instruction must “consider the acquisition of these distinct reading skills and the importance of increasing both the number of words in a student’s vocabulary and the extent of word knowledge for these words” (Ouellette, 2006).

Vocabulary and Reading-related skills of English Language Learners

For English Language Learner (ELL) children, it is important to use reading in English to target breadth and depth of vocabulary. However, there is a necessity to provide some of this reading input in the child’s first-learned language (L1) (Leacox & Jackson, 2012). In fact, in Spanish-speaking bilingual children (where Spanish is L1), it has been demonstrated that the underlying knowledge of lexical items in a child’s L1 helps facilitate acquisition of the word in L2 (in this case English - Lugo-Neris, Jackson & Goldstein, 2010). This idea is reflected in a concept developed by Lugo-Veris et al., (2010) called “vocabulary-bridging”. Vocabulary bridging is the process of defining vocabulary words in L1 in order to expand both breadth and depth of the word in L2. It should be noted that ELL children have been found to be more impaired in depth of word knowledge than breadth even for frequently occurring words (August, Carlo, Dressler & Snow, 2005). However, the breadth of vocabulary may also be limited when compared to English-speaking peers of the same age. Thus, there is a gap in both breadth and depth of vocabulary reflected in ELL children in comparison to other typically developing monolingual children of the same age and grade (August, Carlo, Dressler & Snow, 2005).

A study conducted by Molly F. Collins, explains the effect of rich explanation on Portuguese-English children’s target vocabulary learning through story book reading over a three week period (Collins, 2010). 80 typically developing four and five-year old native Portuguese speakers of varying SES participated in the study. They were all identified as

non-readers by their teachers. Initially, all participants were administered the PPVT III in English and a translation of Form A of the PPVT III. In English the scores ranged from 40-129. In Portuguese, the scores fell between 7 and 63. The intervention itself was conducted solely in English, yet home practice was given in both L1 and L2. Expansion of target vocabulary in this study included the following techniques:

1. pointing to the illustration of the target word
2. providing a general definition of the word
3. providing a synonym
4. making a gesture of the word, when applicable
5. using the word in a context different from that of the book

Target words in this study were unfamiliar English vocabulary words substituted for the more common words in a preschool-level book. Verbs, nouns, and adjectives were targeted with no cognates between Portuguese and English. The children were tested after the reading sessions using a target vocabulary test based on the model of the PPVT-III. The results of the study suggested that the more one reads, the more developed L2 becomes which thus “affects sophisticated reading” (Collins, 2010). It was also shown that language in which the home reading was conducted did not significantly affect sophisticated English language learning. Perhaps most significant was the overwhelming effect of rich explanation to target word learning. Supporting reading with enriched explanation was the strongest contributor to an increase in vocabulary.

PURPOSE

When vocabulary knowledge is impaired treatment will be guided by the hypothesis of vocabulary acquisition favored by the clinician or teacher. The existence of

different hypotheses for vocabulary acquisition has led to the development of varying interventions used to target vocabulary development in children. The purpose of this systematic review will be to **(a)** identify which interventions have been used over the last decade for targeting vocabulary development in Second language learners (especially in second language learners where the L1 was typically Spanish), **(b)** to evaluate the methodological rigor in which those interventions were investigated, and **(c)** determine the clinical relevance of the outcome measurements used to judge if the intervention(s) were efficacious. To accomplish this, I will search for studies published between 2004 and 2014 which utilized some form of intervention to target vocabulary development in groups of children between the ages of 3 to 12 years of age. The overarching aim of this study will be to better inform clinical practice by identifying those interventions which have been shown to demonstrate positive outcomes within research studies that utilized a strong methodological framework, and thus best informing evidence-based practice.

METHODS

Search Strategies

In order to facilitate the search for literature about the best methods for targeting vocabulary treatment in ELL, the following questions were generated to guide the systematic review: **(1)**What methods of treatment have been used to target vocabulary development in Second Language Learning children between the ages of 3 to 12 using group designs?; **(2)** What level of evidence would the methodological structure of those studies be categorized by?; and **(3)** How large are the calculated effect sizes in those studies where data is sufficient for calculation? Electronic databases were searched with this question in mind and with “vocabulary” or “vocabulario” as the main concept in a

keyword search. In conjunction with this concept, the Boolean operators “AND” and “OR” were used to add other terms such as “intervention” (intervencion), “development” (desarrollo), “treatment” (tratamiento), and “program” (programa). Electronic databases of medical and communication sciences were searched including: Pubmed.com, Communication Sciences and Disorders Dome, EBSCOHost, Academic OneFind, and JSTOR. In addition, some references were found from browsing the Journal of Educational Psychology from 2006 and references from Language Disorder textbooks. Since research on methods for vocabulary growth is fairly extensive, parameters in regard to the date range of electronic databases were set for January 2004 to January 2014 (a time span of 10 years). The searches were limited to references in both the Spanish and English language.

Filtering of Studies

Initially references for inclusion were selected based on an analysis of relevant topics (keywords, or descriptions of the participant populations and intervention methods) in the abstract. Only manuscripts utilizing group designs, where statistical analyses were applied to groups of children rather than individual children, were included in the review. The references were further scanned for relevant information about intervention in regards to the facilitation of vocabulary growth in Second Language Learners. Papers were included in the review if they contained (a) the main topics of delayed vocabulary breadth or depth in Second Language Learners, (b) a suggestion for a process in how to treat this delay through intervention, and (c) reported data which would allow some level of ad-hoc analysis.

“Levels of evidence” in regard to different studies were also considered and evaluated based on the American Speech and Hearing Association’s (ASHA) recommended guidelines, as illustrated in Table 1. In addition, it should also be noted that although other review papers are referenced, this paper aimed to review primary sources of treatment information. Of the six levels of evidence defined by ASHA, only levels IIb and above were considered for analysis (IIa, IIb, Ia, and Ib). As such, the papers were limited to controlled studies, quasi-experimental studies, non-experimental studies, and documented clinical experiences. Papers that aimed to answer questions about reading difficulties were excluded. However, it should be noted that both bilingual intervention approaches and non-bilingual approaches were included to demonstrate the process of learning vocabulary in typically developing or monolingual children. The researcher felt compelled to include these studies to analyze the potential of using certain treatments with the bilingual population and speculate on the effectiveness they might have.

Table 1. Levels of evidence used for inclusion criteria of this systematic review.

| Level | Description |
|--------------|---|
| Ia | Well-designed meta-analysis of >1 randomized controlled trial |
| Ib | Well-designed randomized controlled study |
| IIa | Well-designed controlled study without randomization |
| IIb | Well-designed quasi-experimental study |
| III | Well-designed non-experimental studies, i.e., correlational and case studies |
| IV | Expert committee report, consensus conference, clinical experience of respected authorities |

Where possible, data were subjected to an ad-hoc effect-size calculation to determine the clinical significance of reported findings. Effect sizes were calculated using Cohen's d formula, as follows:

$$d = \frac{m_1 - m_2}{S_{pooled}}$$

Where m_1 = pre-intervention mean scores, m_2 = post intervention mean scores, and S_{pooled} = the pooled standard deviations across pre- and post-treatment data.

RESULTS

Published studies that met the criteria for inclusion detailed various treatment strategies ranging from reading intervention to visual cues. Initially there were 11 studies that met search criteria. However, only six studies fell within the time frame of the last ten years and five utilized methodological levels IIb and above. Thus five studies were included in the review. The characteristics of these studies, including the participants, methods, and results, are summarized in Table 2. Unfortunately, due to a lack of relevant studies in Spanish, only English results were included in this review.

Interventions for Targeting Vocabulary Development

The first research question asked was "What methods of treatment have been used to target vocabulary development, especially in Second Language Learning children, between the ages of 3 to 12 using group designs?" Table 2a summarizes the intervention approaches used in the five included studies, while Tables 2b-c identify additional study characteristics. Inspection of the included studies revealed that the majority of participating children came from diverse backgrounds, were learning a second language

(typically English), and ranged from four years of age to six years of age. The bilingual populations within the studies were either children from Spanish-speaking families, Quechua-speaking families, Russian-speaking families, or Portuguese-speaking families.

Table 2a. Characteristics of interventions for included studies.

| Study | Intervention | Control | Setting | Length of Intervention |
|-------------------------------------|--|---|--|---|
| Neugebauer & Currie-Rubin (2009) | 29 Students; 72% home language Spanish; Use of read-aloud (including literal questioning, vocabulary probes, group interaction) | Teachers given same materials but not given explicit instruction on how to use read-aloud to target vocabulary | 4 Bilingual classrooms at a school | 30 minute sessions, 5x per week, 3 weeks |
| Mila Schwartz (2013) | First Language First Model of intervention; bilingual Russian/Hebrew children; does first language first (late introduction of a second language) support progress in depth of vocabulary knowledge in L2 | The main control group was the monolingual L2 preschool group. Control was also present in outside factors such as parental involvement in language practice, gender, age, general short-term memory capability | Bilingual and monolingual preschools in northern Israel but in a separate quiet room | Intervention was every day during normal preschool hours for the duration of the school year with measurements at the beginning and the end |
| Vuattoux Japel, Dion, Dupere (2013) | 22 educators with pre-school children, looking at an educator-implemented intervention; targeting specific vocabulary needs. | Control group without training and intent-to-treat | Non-profit childcare center | 4x per week, 4 months |
| Ruston & Schwanenflugel (2010) | Goal is to determine effectiveness of conversation intervention on expressive vocabulary growth ; pre-kindergarten children receiving intensive conversation with adult | Control group without intensive directed conversation | Preschool center - separate classroom ; control in normal classroom | 25 minutes, 2x per week, 10 weeks |
| Molly Collins (2010) | Effect of rich explanation on vocabulary through story book reading; 4 and 5 year old Portuguese ELLs | Participants matched on L2 receptive scores. One member to experimental and other to control group | Quiet room aside from classroom | 30 minutes, 1x/week, each book for 3 weeks |

Table 2b. Dependent variables (outcome measures) investigated in each included study.

| Study | Outcome Measures | | Description |
|--------------------------------------|--|--|---|
| | Research Created/Criterion Reference | Standardized Test | |
| Neugebauer & Currie-Rubin (2009) | Vocabulario Sobre Dibujos and Comprehension de Textos : Subtests from Woodcock Language Proficiency Batter - Revised (1991) ; Researcher-Designed Curriculum Vocabulary Assessment | Yes | Initially, slight differences in Baseline CVA mean and comprehension mean, extreme in VSD vocabulary mean. Post hoc comparison: significantly better performance for the intervention groups (30% more items correct) |
| Mila Schwartz (2013) | Raven's Colored Matrices (Raven, Raven, & Court, 1998, sets A, B and C); Digit Span (Kaufman & Kaufman 2004); Semantic fluency (Hebrew version adapted from Spreen & Strauss, 1998); Word description (adapted from Rom & Moreg, 1999 and Verhallen & Schoonen, 1993, based on Richard & Hanner, 1985) | Yes | Those using the First Language First model caught up in vocabulary depth to those who had been enrolled in monolingual Hebrew schools in only 2 years of L2 instruction (so L2 develops from advanced level of L1). Note: Those in monolingual schools had L2 immersion at L1 and had lower scores on depth of L1 vocab than L2. Those in bilingual had gradual immersion starting at age 3 |
| Vuattoux, Japel, Dion, Dupere (2013) | Shortened version of Préfontaine and Préfontaine (1968); Dunn et al. 1993; Trudeau et al. 1999 used to find common words | Yes (used to eliminate commonly known words) | Reliable implementation had greater vocabulary gains; intent-to-treat was more reliable; Specific expressive posttest is significantly better in the intervention group. |
| Ruston & Schwanenflugel (2010) | EVT; language sample following Westerveld, Gillon, and Miller (2004) protocol | Yes | Children in intervention group showed greater growth on EVT than controls. Children in intervention group with low vocabulary also showed greater growth in lexical diversity than controls. |
| Molly F. Collins (2010) | Peabody Picture Vocabulary Test-III (Dunn & Dunn, 1997) | Yes | Significant effects found for read-aloud context especially in regards to rich explanation and home reading |

Table 2c. Experimental design and participant characteristics of included studies.

| Study | Design | Description of Participants | |
|--------------------------------------|---|---|---|
| | | Inclusionary Criteria | Total Number and Group |
| Neugebauer & Currie-Rubin (2009) | Quasi-experimental Pretest-posttest control/experimental group (Level IIb) | First Graders in Public School | 55 - 31 male, 24 female; Four groups: Two intervention (29)- Baseline (VSD .47; SD .53) Two Control (26)- Baseline (VSD .9 ; SD .64) |
| Mila Schwartz (2013) | Quasi-experimental Non-equivalent groups (Level IIb) | Second generation Russian/Hebrew speaking immigrants born in Israel; Russian is first and dominant language; Hebrew is non-dominant language; no developmental delay | 31 bilinguals from a bilingual Russian-Hebrew preschool and 20 bilinguals from a monolingual L2 Hebrew speaking preschool without L1 support. |
| Vuattoux, Japel, Dion, Dupere (2013) | Randomized pretest-posttest control/experimental group (Level Ib) | Low income ; preschool (mean age 4 years); attend a child care center | 222 - 46.8% female; 22 groups - 12 intervention, 10 control ; 1 educator per group |
| Ruston & Schwanenflugel (2010) | Randomized Pretest-posttest control/experimental group (Level Ib) | Attending lottery-funded, center-based, full-day prekindergarten programs; any income; parent consent; score falling within normal range of EVT, no prior remediation | 73 - 30 girls 43 boys; six classrooms ; diverse but largely middle class |
| Molly F. Collins (2010) | Randomized pretest-posttest control/experimental group (Level Ib) | Middle-low income; immigrant families from Portuguese-speaking countries; ELL non-readers | Experimental, Control, and No story groups; 80 children (38 girls, 42 boys) |

The aim of a majority of the reviewed studies was to increase both expressive and receptive vocabulary in young second language learners. Each given treatment to achieve this goal differed from the others in at least one significant way. The specific treatment methodologies utilized across the five studies included the following:

- Using the “First language first” model of intervention (Schwartz, 2013)
- Looking at an educator-implemented intervention (Vuattoux, Japel, Dion, Dupere, 2013)
- Learning through conversation (Ruston & Schwanenflugel, 2010)
- Incorporating rich explanation through storybook reading (Collins, 2010)
- Read-aloud (Neugebauer & Currie-Rubin, 2009)

Details of the intervention approaches utilized in the individual studies are further described below.

First Language First Model (Schwartz, 2013)

The First Language First model of intervention involves focusing on L1 for the first few years of a child’s life and then gradually introducing L2. By using this method, second language learners should develop L2 vocabulary with the aid of advanced L1 vocabulary. In this study, 31 bilingual Russian-Hebrew children from a bilingual preschool and 20 bilingual Russian-Hebrew Children from an L2 monolingual Hebrew preschool were assigned to control (monolingual) and experimental (bilingual) group. This study used a non-equivalent groups quasi-experimental design which corresponds to level IIb.

All children represented were born in Israel and their parents considered Russian to be the child’s dominant language. The experiment took place in preschool classrooms

for the duration of a school year. Testing was conducted for 40 minutes at the beginning of the year and 40 minutes at the end of the year. The main goal of this study was to understand whether the First Language First method (late introduction of a second language) supported progress in depth of vocabulary knowledge in L2 in comparison to early immersion models.

After initial testing of semantic fluency and word descriptions (vocabulary depth) through use of Hebrew and Russian adapted standardized tests, it was determined that the bilingual group at the beginning of the year had more depth in L1, while the monolingual group experienced more vocabulary depth in L2 than L1. This was likely because home environment was not sufficient for “paradigmatic and syntagmatic knowledge development” in L1. Compared to one another, the bilingual group had more depth than the monolingual in L1 but less depth in L2. After one year of monitored bilingual education, significant gains were made by the bilingual preschool group in L2. In fact, the children in this group performed at a similar level to their peers in the monolingual group in terms of depth of vocabulary of L2 and performed better in terms of vocabulary depth of L1. Thus, there was a positive effect of later immersion in L2 and an increased rate of L2 vocabulary acquisition.

This outcome implies that the typically accepted “the younger, the better” approach to L2 acquisition may be a misconception and later immersion of second language learners may result in an increased ability to develop vocabulary depth in both L1 and L2.

Educator-Implemented Intervention (Vuattoux, Japel, Dion, Dupere, 2013)

Although an educator-implemented intervention can be used in conjunction with a variety of other models mentioned in this review, it was nevertheless important to examine the effectiveness. In this method, 22 preschool educators from Quebec, Canada (French-speaking) were given the common goal of targeting specific vocabulary needs according to the population they worked with. This occurred at a non-profit childcare center. In this study, some bilingual preschoolers were present but the study was focused on the educator techniques regardless of student language backgrounds. Thus, this study will be examined as a monolingual French language study.

The children that participated in this study had to be preschool age and attending a child care center. From those requirements, there were 222 child participants (46.8% female) who were divided into 22 groups. 12 of these groups were intervention and 10 groups were control with one educator present in each. All of the educators were given the same story books to read to the students. The educators within the intervention group were trained to conduct simulation activities and read the specifically developed storybooks. Each activity was explained and demonstrated and then the educators had the chance to practice in a half-day workshop. The educators were then asked to conduct four 15 minute sessions per week for four months. Pre- and post-tests were conducted in September and February, respectively.

The pre- and post-tests assessed general receptive vocabulary and specific expressive and receptive vocabulary using the French version of the Peabody Picture Vocabulary Test (PPVT; Dunn et al, 1993) and researcher developed measures where the content was aligned with the intervention. The pre-test showed that the children as a whole had comparable general receptive vocabulary and similarly low specific expressive

vocabulary. The post-test showed most significant gains in the specific expressive category.

Analysis of these pre-test results with the post-intervention results shows that intent-to-treat and educator-implemented intervention was highly effective despite the fact that some educators were found to not have implemented the program correctly.

Conversation Intervention (Ruston & Schwanenflugel, 2010)

The conversation intervention sought to determine the effectiveness of expressive vocabulary growth through a conversation in monolingual English-speaking children. As mentioned above, it remains important to analyze as a potential method for expanding vocabulary in bilinguals. The goal of this study was accomplished through intensive conversation between prekindergarten children and an undergraduate senior or graduate student for 25 minutes twice a week for the duration of 10 weeks. 73 prekindergarten children (43 male) were chosen to participate in this intervention based on the criteria of attending full day prekindergarten programs, having parent consent, falling within the normal range of the Expressive Vocabulary Test (EVT), and having had no prior intervention. The design of the study fell under level Ib—a randomized pretest-posttest control and experiment type.

The undergraduate and graduate “talking buddies” were trained for four hours before intervention on good conversational techniques with children and techniques to aid in vocabulary development. For the latter, talking buddies were trained to introduce vocabulary naturally in conversation through “vocabulary recasting and use of rare words”. They were also taught to expand utterances and elaborate on the speech. The

talking buddies practiced the different techniques for 2 hours with different pilot children and were provided with feedback.

The 73 children were then placed into matched pairs that corresponded to their given EVT scores and individual preschool class. Then, one member of each pair was assigned to the control group while the other fell into the experimental group. Children in the control group stayed in the preschool classroom and received no additional minutes of conversation other than that in their regular classroom. The children in the experimental group met in a quiet area of the center.

Testing for this group of children occurred two weeks before and within two weeks after the intervention. The EVT assessment was used to test expressive vocabulary. A language sample was also collected to record the techniques used by the experimenters and the responses of the children participants. It demonstrated each child's lexical diversity and vocabulary breadth as well as depth. After performing analyses of variance (ANOVA) comparing the pre-and post-test scores, it was determined that the experimental group showed greater growth on the EVT than the control group. However, for the experimental children, the intervention did not improve the use of vocabulary in a language sample. There was also a significant benefit of the intervention for children in the experimental group who began the study with low vocabulary skills as compared to the control group. However, the same gains were not made by the children with typically developing vocabulary. Thus, the results of the study show support for an intensive-conversation pullout method in improving vocabulary level of children—especially children with low vocabulary levels.

Rich Explanation through Storybook Reading (Collins, 2010)

Previous research has shown that storybook reading contributes to improved vocabulary breadth and depth. Mary F. Collins took this a step further by adding rich explanation to her study. This study was conducted with four and five year old Portuguese English-language learners with the goal of expanding “sophisticated” vocabulary. It should also be noted that baseline vocabulary and home reading practices were also examined. The interventionists in this study consisted of six preschool teachers who were given instruction on five criteria used to expand story reading. The intervention occurred in a quiet room aside from the classroom for 30 minute sessions once a week. Each book was focused on for 3 weeks of sessions.

80 children participated in the study (42 boys) based on the criteria of being part of an immigrant family from a Portuguese-speaking country and ELL non-readers. These children were divided into three groups upon the start of intervention—experimental, control, and “no story”. Following language pretests using the Target Vocabulary Test (TVT) based on the model of the PPVT-III (Dunn & Dunn, 1997), children were matched on their L2 receptive scores. Then, one member per pair was randomly assigned to the experimental group and the other to the control group. The no-story group was randomly created from the treatment groups to resemble the experimental and control groups with equivalent distributions of L2 scores.

Pretests and posttests were given before (at the beginning of week one) and after (at the end of week 3) each individual book. To test English and Portuguese receptive vocabulary, the Peabody Picture Vocabulary Test-III (Dunn & Dunn, 1997) was given. To test Portuguese receptive vocabulary, the same test was translated and adapted for the Portuguese language. To understand home reading practices, parents were asked to

complete questionnaires on the frequency of reading per week, types of materials read, topics of discussion, and children's interest. The frequency of reading was shown to be a significant contributor to vocabulary learning. Lastly, a target vocabulary test (TVT) based on the PPVT—III (Dunn & Dunn) was administered to test children's knowledge of target words.

Upon pretesting, there were no significant preexisting differences between the groups. After posttests, the treatment group made the largest contribution to target word learning. However, home reading frequency (when in English) and L2 receptive scores also made significant contributions. This was likely because home reading influences L2 and more developed L2 affects a more sophisticated ability to learn new vocabulary. Interestingly, the L1 receptive score did not significantly contribute. It is for these reasons that rich explanation in conjunction with a read-aloud context contributes significantly to target word learning.

Read-Aloud (Neugebauer & Currie-Rubin, 2009)

The methodology for the read-aloud program was to incorporate read-aloud strategies in a discussion-based format to aid the literacy skills and vocabulary knowledge of bilingual students. In this study four teachers were given materials for the read-alouds. Two of the teachers were given explicit instructions on how to use read-alouds to target vocabulary (for the experimental group) while the other two were not given any instruction (control group). This technique was used in 30 minute sessions, five times a week for three weeks in a bilingual (Spanish or Quechua) classroom. This study was performed in Peru and thus the results are specific to a Spanish Language Learner

population instead of English Language Learners, but nevertheless contributes important conclusions for all second language learners.

Fifty-five first graders participated in the study and all were assessed at baseline using the Vocabulario Sobre Dibujos (VSD) and the Comprehension de Textos (CT) subtests from the Woodcock Language Proficiency Battery—revised (1991). The effectiveness of the program was determined by a curriculum-specific vocabulary measure developed by a researcher called the Curriculum Vocabulary Assessment (CVA) which included ten vocabulary items and two practice items that were appropriately challenging but also matched the children’s ability. The students in both the control and the experimental group scored comparably on the CVA at the beginning of the intervention although the students in the intervention group scored lower (on average) on both standardized subtests.

The “Read-aloud pedagogy” that two of the classroom teachers utilized incorporated inferential and literal questioning, vocabulary probes, group discussion, and specific seat arrangements (in a semicircle around the teacher and the book). The other two classrooms had the same materials but did not focus on book-specific vocabulary or a read-aloud format. Instead, these books were read by the students during silent reading and then were retold to the class. After intervention, the CVA was administered again. The results showed a significant difference in the two change scores and students in the intervention group answered 30% more items correctly than the students in the control group with the same baseline score. The VSD indicated a significant change as well in which the intervention group performed higher by 30% again on the post intervention vocabulary measure—even though the baseline scores were higher in the control.

Quality of Study Designs

The second research question asked “What level of evidence would the methodological structure of those studies be categorized by?” The level of evidence specific to each study is identified in Table 2c. The majority of the studies used a level Ib approach which corresponds to a well-designed randomized control study. This study design allows a researcher to comprehend and assess the effectiveness of a given experimental group in comparison to a control group. It also includes both pretests and posttests which allow the researcher to view concrete changes in the child’s vocabulary.

Efficacy of Intervention

The third research question asked “What is the clinical relevance of the outcome measurements used to judge if the intervention(s) were efficacious?” Table 3 summarizes Cohen’s *d* (the effect-size based on comparison of pre-treatment and post-treatment data) calculated from each study. Data from the Neugebauer study was omitted as this study only provided means and standard deviations for the pre-tests. The study did however include statistical significance and a percentage indicating how much greater the experimental group preformed than the control group. This will be considered in the discussion of efficacy. Similarly, in the Collins study, all control data was omitted as well as data for the experimental L2 expressive score. This is due to lack of means and standard deviations for the post-test as well; however some of the effect-sizes were calculated by Collins and included in the results section of the study. She, like Neugebauer, included statistical significance in the results as well so this will also be considered in judging efficacy.

Table 3. Calculated effect sizes for experimental and control groups.

| Study | Tests | Cohen's d Experimental | Cohen's d Control |
|---|-------------------------------|-----------------------------------|------------------------------|
| Neugebauer & Currie-Rubin (2009) | VSD (vocabulary) | N/A | N/A |
| | CT (comprehension) | N/A | N/A |
| | CVA | N/A | N/A |
| Mila Schwartz (2013) | Vocabulary Depth (Russian) | 0.73 | 0.075 |
| | Vocabulary Depth (Hebrew) | 0.52 | 0.44 |
| Vuattoux, Japel, Dion, Dupere (2013) | Specific Expressive | 3.22 | 1.57 |
| | | | |
| Ruston & Schwanenfluge I (2010) | Expressive Vocabulary Test | 0.52 | 0.101 |
| | Lexical Diversity | 0.42 | 0.26 |
| Molly F. Collins (2010) | L1 Receptive Score | 0 | N/A |
| | L2 Receptive Score | 1.15 | N/A |
| | L2 Expressive Score | N/A | N/A |
| | Home Reading/Week | 1.39 | N/A |

Cohen (1988) defined effect sizes as “small, $d = .2$ ”, “medium, $d = .5$ ”, and “large, $d = .8$ ” with the notion that there is always a risk in offering “conventional operational definitions for those terms for use in power analysis in as diverse a field of inquiry as behavioral science”. Thus, while these terms can help define efficacy, they should not be considered absolute.

Using Cohen’s parameters, the smallest effect size in an experimental group across the four studies with data was in the L1 receptive score in the Collins study. However, this did not indicate that intervention did not work—rather that because

intervention was not focused on improving the vocabulary of L1, the development in that lexicon was negligible. There was also a small to medium effect size in judging the lexical diversity of the experimental group in the Ruston study. It should be noted that with the exception of the Vuattoux, Japel, Dion, Dupere (2013) study, all other effect sizes for control groups were small. This indicates that clinically meaningful changes were present in all experimental groups.

Three studies were deemed to have medium effect sizes. These included the EVT of Ruston, and both Russian and Hebrew depth in the Schwartz study. The largest effect was Vocabulary Depth in Russian for the experimental group, which was likely due to the fact that Russian vocabulary depth was continuously emphasized in the bilingual school despite being L1. It should be noted however that the vocabulary intervention was far more effective in the bilingual school (experimental group) compared to the monolingual school (control group), characterized by the experimental group manifesting two less years of language learning in L2 than the control group and yet was still able to succeed them in vocabulary growth.

Finally, two studies indicated large effect sizes. The specific expressive category of the Vuattoux study far exceeded any other effect sizes with a calculation of $d = 3.22$. This indicates that it had the least overlap between pre-treatment and post-treatment measures. The Collins study was consistent in exhibiting large effect sizes for both the L2 receptive score and the home reading. However, without the control group effect sizes, it is difficult to compare to determine accurate efficacy. Nevertheless, Collins makes the claim that all three categories contributed significantly to vocabulary depth and breadth.

Discussion

A range of techniques have been used to facilitate the growth of vocabulary in both monolingual and bilingual populations. The common aim of these treatment strategies was to increase breadth and depth as well as receptive and expressive vocabulary—especially with emphasis on second language learners. The bilingual treatment methods included read-alouds, rich explanation, and the first language first model. The monolingual treatment methods included an educator-implemented approach and a conversation-based approach and were included in this review in order to provide possible treatments for vocabulary developments in second language learners.

The most effective treatment in the bilingual category focused on rich explanation of story book reading—the study by Collins. This had the second largest effect size of all the studies and showed drastic improvement in the L2 receptive scores and home reading. However, it should be noted that due to lack of means and standard deviations for post-tests, the efficacy of this study was partly judged by the claims of Collins in her results section (stating that all three factors contributed significantly to a drastic increase in the vocabulary of bilingual students). The read-aloud technique (Neugebauer & Currie-Rubin, 2009) could also be deemed as a highly effective technique as the students in the experimental group were said to have performed 30% better than the students in the control group on both the CVA and the VSD (vocabulary). Thus, the least significant of the three in terms of mere statistical analysis was the first-language-first model. However, it should be noted that significant gains were made by the experimental group over the control group despite the control group having more years of second language use. Thus, it could be argued that this was the most effective in terms of quick vocabulary acquisition while maintaining the depth of the first language.

For the monolingual treatment studies, one stood out far above the other and had the highest effect-size of all studies. This was in reference to the educator-implemented with intent to treat study (Vuattoux, Japel, Dion, Dupere, 2013). This technique was likely the most effective because it used a variety of different methods to target vocabulary through in-class learning with other peers. The conversation intervention by Ruston & Schwanenflugel (2010) was a monolingual study and had the least effect of all the studies and a far lesser effect on expressive vocabulary as compared to the educator-implemented strategies.

SUMMARY

The purpose of this systematic review was to (a) identify which interventions have been used over the last decade for targeting vocabulary development in second language learners where the L1 was typically Spanish, (b) to evaluate the methodological rigor in which those interventions were investigated, and (c) determine the clinical relevance of the outcome measurements used to judge if the intervention(s) were efficacious. Results indicated that a variety of interventions have been used to target vocabulary over the past decade, although only 5 studies could be characterized as using strong research designs, which were predominately level IIb designs. Based on reported effect sizes, the most effective method for developing the vocabulary (breadth and depth) of a second language learner is best accomplished through storybook reading with rich explanation especially through educator implementation and home reading in the second language.

One limitation of this review is that it only looked at published findings which may have resulted in a bias towards positive interventions. In other words, no studies

were found that discourage the use of certain techniques. However, it remains useful to examine the most recent strategies and their effectiveness in order to establish new research on vocabulary intervention as well as structure current intervention in the more effective way possible for second language learners.

REFERENCES

- Adams, M.J., & Huggins, A.W.F.(1985).The growth of children’s sight vocabulary: A quick test with educational and theoretical implications. *Reading Research Quarterly*, 20, 262–279.
- August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for english language learners. *Learning Disabilities Research and Practice*, 20(1), 50-57.
- Biemiller, A. (2006). *Vocabulary: The missing link between reading and literacy*. Paper presented as keynote speech at International Reading Association World Congress.
- Brooks, D. (2010, July 09). The medium is the medium . *The New York Times*, p. A23.
- Bus, A.G., Ijzendoorn van, M.H., & Pellegrini, A.D. Joint Book Reading Makes Success in Learning to Read: A Meta-Analysis on Intergenerational Transmission of Literacy. *Review of Educational Research*, 1995: 65(1),1-21.
- Chiappe, P., Chiappe, D., & Gottardo, A. (2004). Vocabulary, context, and speech perception among good and poor readers. *Educational Psychology*, 24(6), 825-843.
- Collins, M.C. 2010. ELL preschoolers’ Englishvocabulary acquisition from storybookreading. *Early Childhood Research Quarterly*, 25 (1): 84–97.
- Duferense , T., & Masny, D. (2006). Multiple literacies: Linking the research on bilingualism and biliteracies to the practical. *Paediatr Child Health*, 11(9), 577-579.
- Eldredge, J.L., Quinn, B., & Butterfield, D.D.(1990).Causal relationships between

- phonics, reading comprehension, and vocabulary achievement in the second grade. *Journal of Educational Research*, 83, 201–214.
- Justice, L.M. & Ezell, H.K. (2002). Use of storybook reading to increase print awareness in at-risk children. *American Journal of Speech-Language Pathology*, 11, 17–29.
- Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal*, 73(4), 440-464.
- Lady of the Lake University. (2012, March). *Homeless children with language delay*. Presentation delivered at TSHA Texas speech and hearing association 2011, San Antonio, TX.
- Leacox, L., & Jackson, C. W. (2012). Spanish vocabulary-bridging technology-enhanced instruction for young english language learners. *Journal of Early Childhood Literacy*, 0(0), 1-23. doi: 10.1177/1468798412458518
- Miller, G. E. (1941). Vocabulary building through extensive reading. *The English Journal*, 30(8), 664-666.
- National Center for Education Statistics (2012). The Nation’s Report Card: Vocabulary Results From the 2009 and 2011 NAEP Reading Assessments (NCES 2013 452). Institute of Education Sciences, U.S. Department of Education, Washington, D.C.
- Neugebauer, S. R., & Currie-Rubin, R. (2009). Read-alouds in Calca, Peru: A bilingual indigenous context. *The Reading Teacher*, 62(5), 396-405.
- Lugo-Neris, M. J., Jackson, C. W., & Goldstein, H. (2010). Facilitating vocabulary acquisition of young english language learners. *Language, Speech, and Hearing Services in Schools*, 41, 314-327.
- Owens, R. E. (2010). *Language disorders: A functional approach to assessment and*

- intervention*. (5th ed., pp. 377-385). Boston, MA: Pearson Education, Inc.
- Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Education Psychology*, 98(3), 554-566.
- Ruston, H. P., & Schwanenflugel, P. J. (2010). Effects of a conversation intervention on the expressive vocabulary development of prekindergarten children. *Language, Speech, and Hearing Services in Schools*, 41, 303-313.
- Shen, Z. (2008). The roles of depth and breadth of vocabulary knowledge in efl reading performance. *Asian Social Science*, 4(12), 135-137.
- Schwartz, M. (2013). The impact of the first language first model on vocabulary development among preschool bilingual children. *Reading and Writing*, 27(4), 709-732. doi: 10.1007/s11145-013-9463-2
- Vuattoux, D., Japel, C., Dion, E., & Dupere, V. (2013). *Targeting the specific vocabulary needs of at-risk preschoolers: A randomized study of the effectiveness of an educator-implemented intervention*. Département d'éducation et formation spécialisées, Université du Québec à Montréal, Montreal, QC, Quebec, Canada., Available from PubMed. (10.1007/s11121-013-0379-5).
- West, M. R. (2012). *Is retaining students in the early grades self-defeating?*. Informally published manuscript, Education, Brookings Institute, Washington, DC, .
- Woodcock, R.W.(1998). WRMT-R: Woodcock Reading Mastery Tests—Revised. Circle Pines, MN: American Guidance Services