Examining the Predictors of Response to Treatment for Children with Primary Language Delays

By

Lauren Hazledine Hampton

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

In partial fulfillment of the requirements

For the degree of

DOCTOR OF PHILOSOPHY

in

Special Education

May 2016

Nashville, Tennessee

Approved:

Ann P. Kaiser, Ph.D.

Stephen Camerata, Ph.D.

Mary Louise Hemmeter, Ph.D.

Megan Y. Roberts, Ph.D.

Sonya Sterba, Ph.D.

To my endlessly supportive husband and darling baby girl

ACKNOWLEDGEMENTS

My work during my time at Vanderbilt could not have been possible without the support I received from the Institutes of Educational Sciences (H325D0100034 and R324A090181) and the Vanderbilt Semmel Award.

I have had the opportunity to work with the most remarkable researchers who have provided me with opportunities and guidance throughout my program that will forever shape who I will be as a researcher. I have the greatest appreciation for my advisor, Dr. Kaiser, who has provided me with constant support, endless opportunities around the world, and who has stretched my thinking in every dimension. My wise dissertation committee has allowed me to develop thought-provoking questions and exciting analyses through their expertise and guidance. Additionally, this work would not have been possible without the wisdom and leadership that Dr. Roberts provided.

I would like to express my sincere thanks to the amazing KidTalk research team (Elizabeth Fuller, Jodi Heildage, Nicole Dhurv, Kimberly McCulla, Suzanne Thompson, Stephanie Jordan, and Sabrina Smiley Evans), who have enabled me to ask and pursue the most interesting intervention research questions. I am grateful for the endless enlightening discussions and the positivity that has allowed me to be creative and have fun over the last four years.

Finally, I'd like to thank my family. To my mother: you have cheered me on and allowed me to become an inquisitive thinker. To my father: you dreamed greater for me than I have ever allowed myself to dream. My incredible husband: your support through late nights and early mornings with an optimistic attitude allowed me to stay sane throughout graduate school while having a baby. To my entire family: your love and support from around the globe have fueled my journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ASD | Autism Spectrum Disorder (ASD)

NDW | Number of Different Words (NDW)

PLS | Preschool Language Scale-4th edition (PLS)

T0 | Time point 0, before intervention (T0)

T1 | Time point 1, 1-month into intervention (T1)

T2 | Time point 2, 2-months into intervention (T2)

T3 | Time point 3, post-intervention, 3 months from baseline (T3)

T4 | Time point 4, 6 months following intervention (T4)

T5 | Time point 5, 12 months following intervention

Bayley | Bayley Scales of Infant Development (Bayley, 1993)

RMSEA| Root Mean Square Estimate of Approximation

AIC | Akaike Information Criteria

EMT | Enhanced Milieu Teaching

CCX | Caregiver-child interaction

STAT | Screening Tool for Autism in Toddlers and Young Children (Stone & Ousley, 1997)

SPA | Structured Play Assessment (Ungerer & Sigman, 1981)

LGM | Latent Growth Curve Model

CHAPTER I

INTRODUCTION

As many as 17% of young children present with language delays of no known etiology (Leonard, 2014; Tomblin et al., 1997). Primary language delays are not associated with deficits in cognition, sensory disorders, or other developmental, medical or genetic diagnoses. Although some of these children may catch up to their typically developing peers' language abilities, a considerable number of children will demonstrate persistent, long-term language-related deficits in academic and social skills. These long-term effects may include difficulties in developing a strong working memory, complex vocabulary, advanced social skills or robust reading comprehension skills (Leonard, 2014; McCormack, McLeod, McAllister, & Harrison, 2009).

Within the population of young children with primary language delays, some children catch-up to their typically developing peers and appear to have no further language delays once they reach kindergarten (Leonard, 2014). Due to this apparent spontaneous recovery, there is a tradition of delaying intervention with this population until it is clear that the delays are persistent (Paul, 1993, 2000). Children with expressive language delays and typical receptive language, appear to recover at a greater rate than children with expressive and receptive delays (Buschmann et al., 2008). Recovery, however, is poorly understood and the child characteristics associated with positive outcomes in this population require further examination.

Not only is the observed " recovery" in some children with primary language delays poorly understood, but also there is also little information about long-term outcomes of early intervention and the moderators of treatment outcomes within this population (Nelson, Nygren, Walker, & Panoscha, 2006). Understanding the moderators of treatment outcomes is important for individualizing treatments and for better understanding which children within this population require early intervention. More accurate early identification of children who would benefit from treatment potentially could reduce societal costs and family stress (Paul, 2000).

There is considerable heterogeneity among young children with primary language impairment. Early language delays vary and may be deficits in expressive language only, in receptive language only, or in both receptive and expressive language. Additionally, children may present with delays in other skills related to language development including speech and joint attention skills (Bishop & Leonard, 2014; Leyfer, Tager-Flusberg, Dowd, Tomblin, & Folstein, 2008; Paul & Jennings, 1992; Paul & Shiffer, 1991). In young children, early language delay can be an indicator of several potential diagnoses. For example, there is overlap between early significant language delays and autism spectrum disorder. As many as 40% of children first identified with language impairment ultimately meet the Autism Spectrum Disorder (ASD) criteria for delays in social-communication, despite the fact that a language delay alone does not imply a social-communication delay (Leyfer et al., 2008).

Additional areas of developmental concern may be identified when children with early language delays reach kindergarten age. As many as 65% of children first identified with early language delays will be diagnosed with reading delays at age 8 compared to only 5% of their typically developing peers (Leonard, 2014, p. 168). Children with early language delays often have fewer social contacts in the first years of school and their parents and teachers rate their social skills as weaker than their peers (Fujiki, Brinton, Hart, & Fitzgerald, 1999; Fujiki, Brinton, Morgan, & Hart, 1999; Stanton-Chapman, Justice, Skibbe, & Grant, 2007).

Predictors of language development

Examining the predictors of language development in typical children may inform effective intervention and further the identification of predictors of children's response to treatment. Further, examining predictors may suggest factors that moderate the effects of early intervention. In typical development, diversity and frequency of gesture use predict positive language development outcomes (Rowe & Goldin-Meadow, 2009; Watt, Wetherby, & Shumway, 2006). Similarly, early socio-cognition abilities predict superior and more complex social communication abilities throughout development (Chiat & Roy, 2008). Language comprehension in typical toddlers predicts greater expressive language ability by preschool age (Watt et al., 2006). Moreover, diversity in early speech sounds predicts greater expressive language abilities and greater morpho-syntax development (Chiat & Roy, 2008; Watt et al., 2006). Thus, in typical development, important indicators of later strong expressive language abilities include gesture use, receptive language, and diversity in speech sounds.

For children with early language or developmental delays, similar early communication skills predict later language. Gesture attainment and overall communicative rate predict better expressive language development in children with developmental delay (Brady, Marquis, Fleming, & McLean, 2004). Cognitive abilities have not been found to significantly predict language outcomes among children with language impairment who have no other developmental delays, but vocabulary size at age 2 has been identified as an important early indicator related to later expressive language abilities (Bishop & Leonard, 2014; Feldman et al., 2005). Additionally, poorer speech sound diversity at age 3 has been shown to predict additional challenges in academic and social domains as children enter school above and beyond language deficits (McGrath et al., 2008). In general, the currently identified predictors of later language abilities in children with early language impairment are similar to those for children with typical development: gesture use, vocabulary size, and speech sound repertoire.

For children with autism, who may share some early language characteristics with children with primary language impairment, numerous studies have identified key predictors of language outcomes. Although children with autism are likely to be more impacted in the domain of social communication, some features of their language are similar to children with early language delays. Therefore, understanding language development and intervention moderators of language outcomes in this population may provide insight into variables that may be important for children with primary language impairments.

Young children with autism who have early, frequent and diverse joint attention gesture use, strong receptive language skills, and greater speech sound diversity, ultimately develop better expressive language abilities (Bopp & Mirenda, 2011; Deconinck, Soncarrieu, & Dan, 2013; Toth, Munson, Meltzoff, & Dawson, 2006; Yoder, Watson, & Lambert, 2014). For children with autism, imitation ability, object interest, and toy play have also been identified as important significant predictors of spoken language outcomes (Smith, Mirenda, & Zaidman-Zait, 2007; Thurm, Lord, Lee, & Newschaffer, 2006; Toth et al., 2006). Because the language abilities of children with ASD and those of toddlers with language impairment may be similar, it is possible that imitation, object interest, or toy play may also be important predictors of future expressive language abilities for toddlers with language delays (Leyfer et al., 2008).

In sum, evaluating the predictors of language skills observed in children with typical development, children with language delays, and children with ASD could contribute to understanding long-term outcome improvements observed during language interventions for children with language delays. Currently, few moderator analyses have been conducted across these populations, and no analyses have examined moderators of language outcomes with toddlers who have primary language delays only.

Interventions

The literature examining the effects of interventions for young children with language delays is limited, (Buschmann, 2009; Glogowska, Roulstone, Enderby, & Peters, 2000; Law, Kot, & Barnett, 1999; Roberts & Kaiser, 2015; Wake et al., 2011). Studies have demonstrated minimal to moderate short-term effects on proximal and context-bound language measures, but few effects on standardized language assessments (Table 1). Additionally, no studies have identified the characteristics of children for whom interventions are most effective or the intervention conditions (e.g., dosage, choice of implementer, targets of intervention) from which children benefit most. Because there is little evidence indicating which intervention practices result in the best long-term outcomes and studies have not examined specific child characteristics that predict spoken language outcomes, it is difficult to make recommendations to practitioners about how to best intervene with this population.

For example, few effects were observed following a randomized trial of a low-dose parent training procedure for toddlers with primary language delays compared to community services (Law et al., 1999). Effects in this study ranged from d= -0.52 on distal standardized measures of language to d= -0.31 on proximal language sample measures. Parent report of vocabulary indicted moderate effects (d=0.47) for the treatment group;. However, in this parent training study, it is possible that observed effects may have been partially the result of changes in parent’s perception rather than changes in the child’s language ability alone.

In another low-dose study, older preschool children with primary language impairments demonstrated differential gains only in receptive language following eclectic speech and language therapy (average therapy occurred just 6 hours total during the intervention period), as compared to children who were waiting to receive treatment (Glogowska et al., 2000). There was no evidence of improved spoken language outcomes or a reduction in errors in phonology. Glowgowska and colleagues reported that there were few indications of spontaneous recovery in either group suggesting that intervention should not be delayed in this population (2000). However, the lack of spoken language improvements suggests that different intervention strategies, greater dosage or more precise fidelity of intervention may be necessary to achieve optimal spoken language outcomes for this population.

More recently, in a parent-implemented intervention randomized trial of 61 toddlers with expressive language delays (and no other developmental delays or receptive language delays), toddlers in the intervention group demonstrated greater improvements than children on the waitlist (Buschmann, 2009). This study used a 3-month program for parents focusing on responsiveness, language modeling and prompting, and shared book reading. The results indicate that only 8% of the intervention participants remained delayed in expressive language, as compared to 26% of the waitlist group. Two considerations must be made when interpreting these results. First, the sample size is small (only 47 total in the final analysis) and the participants did not present with receptive language delays or any other impairment often associated with early language delays. This study presents some evidence about the effects of parent-implemented interventions for toddlers with expressive-only delays, but does not address the range of children with early expressive and receptive language delays. Secondly, this study does not indicate which participants responded the best to the intervention or provide specific information about the children who did remained delayed.

In a third large (n=301) randomized trial implemented a low-dose parent training ("You Make A Difference") was compared to a business as usual control to address language delays in 18-month old children (Wake et al., 2011). Any child at 18 months of age, who was in the 20th percentile or below for either expressive or receptive domains, was included in this population-based Australian study. There were no significant effects for receptive or expressive language, vocabulary, or problem behaviors, however the low-dose intervention was feasible and acceptable to parents. Although the parent intervention was manualized, the intervention was shortened from the nine prescribed sessions to only six. No measure of fidelity was reported for parent training or the parents' use of the training with their children. More rigorous parent-implemented intervention studies are necessary to better understand the impact of a well-implemented, high-quality parent-delivered intervention on the language development in children with primary language delays.

A recent intervention study for toddlers with receptive and expressive language delays is the basis for the current analysis. Roberts and Kaiser (2015) examined the effectiveness of a clinician and parent co-delivered intervention for toddlers in a randomized trial enrolling 97 toddlers averaging 30 months at the beginning of the study. Toddlers in the intervention group received Enhanced Milieu Teaching (EMT), a naturalistic play-based language intervention delivered by a master's-level interventionist during three weekly intervention sessions that included systematic parent training using the Teach-Model-Coach-Review procedures (Kaiser & Roberts, 2013a; Roberts & Kaiser, 2015). Results immediately following intervention indicated that toddlers who received the EMT intervention package improved in their receptive language abilities assessed on the PLS-4 (Zimmerman, Steiner, & Pond, 2011) and expressive vocabulary diversity (number of different words [NDW] produced in a language sample) as compared to the toddlers in the control group who received community services. Although differences between treatment and control groups in expressive language outcomes as measured by the PLS-4 were not significant immediately following intervention, on average children in the intervention group gained 9 points and children in the control group gained 5 points. The variability for both groups (SD= 13.9,12.0) suggested that some participants did make substantial gains in expressive language while others did not improve during the period between pre and post testing.

Response to treatment

Measuring children's response to treatment while controlling for the growth in the comparison group is important for both research and practice (Yoder & Compton, 2004). Especially within heterogeneous populations such as toddlers with language delays, it is important to determine for whom the intervention is most and least effective. Such information is needed in order to make recommendations to practitioners about tailoring interventions, and to make recommendations about children for whom the intervention is likely to be effective (Warren, Fey, & Yoder, 2007).

The development of sensitive indicators of children's response to language intervention is also important for advancing research methods (Almirall, Collins, & Murphy, 2012). Currently, novel research methods in education are moving towards developing adaptive intervention designs. These designs rely on an early measure of response to treatment to make decisions about sequential treatments that may improve long-term outcomes. The design does not prescribe one static intervention package to a group, but prescribes an adaptive sequence of treatments based on participants' response to the initial treatment (Collins, Murphy, & Strecher, 2007). An early measurement point allows researchers to evaluate individual participant's progress, relative to a pre-determined benchmark, and make adaptations to the treatment model to optimize outcomes (Nahum-Shani et al., 2012). Although adaptive treatment designs are common in drug-intervention research, data- based benchmarking and development of adaptive interventions is a novel development in educational, behavioral, and language interventions (Almirall et al., 2012; Joss et al., 1994; Thall et al., 2007). By understanding who, under what conditions children with language delays are most likely to respond to treatment, we can better tailor adaptive interventions by making adjustments based on pre-test performance. Better yet, if we can understand the early indicators of response to treatment by examining individual trajectories, we can create benchmarks to best tailor adaptive interventions.

Study Purpose

In order to expand upon the current intervention literature for toddlers with primary language impairments, and to begin to contribute evidence towards a model for response-to-treatment with benchmarking for this population, the current study expanded the primary analysis of the data collected during a RCT study of an early intervention for toddlers with receptive and expressive language delays (Roberts & Kaiser, 2015). The current analysis examined predictors of spoken language outcomes for toddlers with primary language impairments in order to determine the characteristics of children for whom intervention was most effective and examined measures of progress during early treatment to determine which measures predicted treatment outcomes.

The following objectives were examined using a latent growth curve modeling approach to estimate and predict the intercepts and slopes, and how growth interacts with group assignment over time, using the predictors outlined in Table 1 (receptive language, speech repertoire, imitation, joint attention and play complexity).

Objective 1. The first objective was to determine which pre-intervention child characteristics interacted with time and group assignment for proximal outcomes, as measured by language diversity (number of different words during a language sample: NDW), and distal outcomes, as measured by the Preschool Language Scale-4th edition (PLS-4 total score) immediately following intervention and 12-months following intervention. Candidate predictor variables were selected on the basis of literature on child characteristics predicting toddlers' language development in typical and language-delayed populations. This first objective was addressed using a latent growth curve model to examine the trajectory of NDW and PLS-4 scores over the assessment time points.

This objective included the following four research questions: 1a) What pre-intervention child characteristics interact with group assignment to predict expressive vocabulary (a proximal measure; NDW) immediately following intervention (T3)? 1b) What pre-intervention child characteristics interact with group assignment to predict expressive vocabulary (a proximal measure; NDW) 12 months following intervention (T5)? 2a) What pre-intervention child characteristics interact with group assignment to predict total language ability (a distal outcome; PLS-4) immediately following intervention (T3)? 2b) What pre-intervention child characteristics interact with group assignment to predict growth in total language ability (distal measure; PLS-4) 12 months following intervention (T5)? The working hypothesis was that toddlers in the intervention group who demonstrated larger speech sound repertoire, higher rates of spontaneous imitation, higher rates of social skills, and greater diversity in play actions at the first assessment point would respond best to the intervention as measured by better performance on proximal and distal outcome measures as compared to the growth in toddlers in the community group immediately following intervention and 1-year following intervention.

Objective 2. The second objective was to examine the early indicators of response to treatment in the treatment group as compared to the language growth in the control group. This objective included two research questions: 3a) Does improvement in expressive vocabulary (NDW) during the first 4-weeks (as measured by the change between T0 and T1) interact with group assignment to predict the total standardized language outcomes (PLS-4) immediately following intervention (T3). 3b) Does early growth (change between T0 and T1) predict performance 1-year following the initial study period (T5)? The working hypothesis is that group assignment will interact with expressive vocabulary growth over time (NDW) to predict greater gains on a standardized distal measure of language at the post-test (T3, PLS-4), and at follow-up (T5).

CHAPTER II

METHOD

Data set

This analysis used the dataset and video database from a randomized trial (NCT01975922) comparing Enhanced Milieu Teaching (EMT) to community business as usual treatments for toddlers with language delays (Roberts & Kaiser, 2015). The dataset included observational, standardized, and parent report measures for 45 participants in the intervention group and 43 participants in the comparison group. Additionally, a sample of 88 typically developing toddlers was included as a comparison sample for normative development. Data were collected at six time points: immediately prior to intervention (T0), 1 month intervals during the 3-month intervention (T1, T2), immediately following the 3-month intervention (T3), 6 months following intervention (T4), and 12 months following intervention (T5) (See Table 3).

Participants

Language delayed toddlers were recruited to participate, and were between 24 and 42 months of age at the beginning of the study (Mean age 30.5 months, standard deviation 5.1 months) and were mostly boys (82%). All of the child participants met criterion for language delays as defined as performance at least 1.3 standard deviations below the norm on a standardized assessment of expressive and/or receptive language (Bayley Scales of Infant Development; Bayley, 1993). Toddlers with a cognitive disability, medical condition, hearing loss, or autism spectrum disorder were excluded from participation. Ninety-seven toddlers and their caregivers consented to participate in the study; 45 participants were randomly assigned to the intervention group and 53 to the control group (see Table 2). Attrition included nine participants in the control group over time. Two of these participants completed testing only during the screening assessment period (T0) and then withdrew from the study. Children who completed all of baseline testing were retained in the growth curve analysis, however the two children who only completed screening assessments were removed due to the large amount of missing data across assessments.

Participants in the control group received usual treatments for speech and language delays, however, in Tennessee children under the age of 3 do not qualify for public early intervention services when they present with delays in language only. Thus, only 20% of children in intervention and 14% of children the comparison group reported receiving any intervention services during the study. Sixty-nine of the participants presented with both receptive and expressive language delays (1.33 SD below norm), and 28 with expressive delays only. Although children with a diagnosis of ASD were excluded from the sample screening, 13 participants were diagnosed with autism spectrum disorder (ASD) at some point during the intervention or follow up periods. Participants with ASD were evenly distributed across experimental groups and the subgroups of participants with expressive-only and expressive and receptive delays. No significant differences between the control group and intervention-group were observed on any observational measures during the initial testing period (T0, Table 2).

An additional typical comparison sample was recruited from the community (pediatrician's offices, parenting magazines, and the university) to participate in the study. Children were on average 29.68 months old (SD=5.86) at the first assessment with average cognition and average language skills (Table 2). Children in the normative sample were assessed at the same six time points as children participating in the intervention study. In the current analysis, the typical children are included as a comparison to benchmark normative development.

Setting and Materials

All assessments occurred in a university clinic setting and were administered by masters' level assessors who were not involved in the intervention for the child or family (Roberts & Kaiser, 2015). Assessors were trained to criterion on administration of all protocols prior to the beginning of the study. The assessment room included a child-sized table and chairs, a play mat on the floor, and open shelves for testing materials. Two-thirds of intervention sessions and the six workshops in which parents were first taught the intervention components occurred in a similar university-based clinic room. The remaining intervention sessions (one per week) occurred in the families' home during every-day routines in rooms selected by the family. Sessions in the clinic included a variety of toys appropriate for the child, a snack, and a selection of books. At home, the parents selected toys, books, a snack and household items typically used by the family. All sessions and assessments were video recorded using a digital camera.

Intervention

The language-delayed toddlers were randomly assigned to receive 28 intervention sessions or business as usual community treatments. The one-hour long intervention sessions occurred three times per week, twice in the clinic and once at home. The Enhanced Milieu Teaching manualized intervention included the Teach-Model-Coach-Review procedures for parent training and direct therapist intervention delivery (Roberts & Kaiser, 2015). During the play-based intervention sessions, interventionists and parents used six main components to teach and facilitate language learning: responsiveness, turn taking, target language modeling, language expansions, time delays, and prompting strategies. During each session, the therapist played with the child for 10 minutes and used all the EMT procedures while the parent observed. Then, the caregiver practiced for 10 minutes in play while the therapist provided coaching and support. The therapist and parent implemented the EMT strategies in a book reading and snack routine, with the parent managing the activities and serving as the primary implementer over time. The child was required to communicate only during milieu teaching prompting episodes that occurred between 5-10 times during each 10-minute interventionist-implemented intervention session. Additionally, the parent prompted the child at the same rate once they had been taught the prompting strategy. Fidelity of the therapist implementation of the EMT sessions was assessed for 20% of sessions and averaged 94% across strategies.

The parent-training component of the study included the Teach-Model-Coach-Review procedures using EMT criterion performance benchmarks to provide the parents with data-based feedback on performance (See Kaiser & Roberts, 2013 for a full description). These procedures ensured that caregivers could implement the six core EMT strategies at fidelity by the end of the intervention period. Parents were trained during workshops that occurred throughout intervention as the parent reached criterion levels for the pervious skill. Workshop content was delivered using video examples, handouts, worksheets, and role-playing with the parent and included examples specifically tailored to the individual child. Parent implementation of the strategies was monitored during intervention sessions. Generalization to caregiver child interactions outside of the intervention sessions was assessed during clinic observations at the primary assessment points (T0-T5). These assessments, administered to both groups, were designed to capture the extent to which the parents used the EMT strategies over the course of intervention as well as the extent to which they maintained the intervention strategies following intervention (Roberts & Kaiser, 2015).

Measures

During each of the six assessment time points in the study, a full battery of standardized, observational, and parent report assessments was completed. The caregiver provided information for demographic characteristics, intervention services received, and any medical information about the child at each assessment point. The following eight constructs were measured and are used in the current analysis (Measures for each time point are summarized in Table 3).

Vocabulary diversity. A 20-minute naturalistic language sample (administered by a trained assessor using five standard toy sets and following a standard protocol) assessed the number of different words a child used during a play interaction with an adult. The assessor engaged with the child and the toys, taking turns in play, but not modeling specific content or target language. The assessor responded to all of the child's communicative attempts by repeating his or her language or using expressives (e.g., wow, uhoh, uh-huh). All language samples were transcribed and coded from video using the Systematic Analysis of Language Transcripts (SALT) software (Miller & Chapman, 2008). The child's expressive vocabulary diversity was measured by estimating the number of different words (NDW) used during the observation. Twenty percent of the language samples were coded for reliability. All transcription and coding was completed using the KidTalk transcription and coding manual (Appendix A). All assessments were video recorded and 20% of all assessments were rated for fidelity of implementation and reliability of scoring was determined for 100% of all assessment protocols. The language sample was conducted at every assessment point before, during, and following intervention (T0-T5). Research assistants blind to the study conditions transcribed language samples. A second research assistant verified each language sample. Analysis of NDW was completed automatically using the SALT program (Miller & Chapman, 2008).

Cognition. Cognitive skills were assessed using the Bayley Scales of Infant Development (Bayley, 1993). The Bayley has a mean score of 100 and standard deviation of 15. The Bayley scale composite score for cognition demonstrates a split-half reliability of .91, and the validity as estimated between .72-.79 for typically developing toddlers. The Bayley was used as a screening measure to ensure that participants did not demonstrate significant cognitive delays and was administered immediately prior to intervention (T0) only.

Receptive language. Receptive language abilities were assessed using the receptive communication subscale of the Bayley. This measure of receptive language was administered only at the first assessment period (T0). Additionally, the Preschool Language Scale, 4th edition (PLS-4; Zimmerman, Steiner, & Pond, 2011) was administered at all time-points and auditory comprehension subscale was used as a separate measure of receptive language, but was not used specifically in this analysis as a predictor due to the high correlation with the Bayley and the redundancy in the PLS-4 growth curve analysis.

Language ability. The overall language ability of each child was also estimated from the PLS-4, which was administered at four time points: prior to intervention (T0), immediately following intervention (T3), and at both follow-up time points (T4 and T5). The PLS-4 total score was used as an outcome measure in the current analysis (T3-5). Overall language ability, as indicated by the PLS-4 total score, is a composite of the expressive communication and auditory comprehension subscales. The total language score has a mean of 100 and standard deviation of 15. Fidelity of administration of the PLS-4 was assessed for 20% of administrations and was rated higher than 95%, and 100% of protocols were scored by two independent raters to ensure correct scoring.

Speech sound repertoire. A measure of the child's speech sound repertoire was coded from the video recordings of the initial language sample (T0). An experienced speech language pathologist with prior training in speech transcription coded each speech sound produced during the 20-minute language sample. All consonant sounds and long vowel sounds were categorized based on sound type; sounds were scored regardless of whether they were used communicatively or not. The number of different speech sounds produced at least once and number of different sounds produced more than once was recorded. Additionally, each of the different forms were rated for approximate a level of complexity. The number of intervals of silence was also recorded. A second speech language pathologist independently coded 20% of the language samples; the samples were randomly selected. When agreement was below 80% for any language sample, agreement and disagreements were discussed then consensus coded by the coding team (Appendix B).

Complexity of play. The highest level of play observed during the Structured Play Assessment was used to estimate complexity of play skill (SPA; Ungerer & Sigman, 1981). The SPA was administered by a trained assessor at the beginning of the study (T0), following intervention (T3), and during both follow up time points (T4 and T5). The SPA includes five toy sets that are presented individually to provide opportunities for the child to demonstrate play skills at four levels (simple play, combination play, pre-symbolic, and symbolic play). The assessor presented and arranged each toy set. She initially played with the toys only in ways the child played with the toys. After the child has explored the toys, the assessor modeled two play actions at a higher level of play than the child demonstrated during initial play. The highest level of play observed during each play set was scored. A total play score was calculated by averaging, the score for highest level of play scored for each of five toy sets to account for differences in opportunities across children and to reduce missing data.

Joint attention. Use of joint attention behaviors (pointing, showing and giving to share) was coded from observations of caregiver-child play interactions (CCX) administered at the beginning of the study. During this assessment, parents were given a standard toy set and were prompted to play with their child for 10-minutes. Parent and child communication during these sessions were transcribed and coded using SALT. Child communication included the use of joint attention behaviors and the total number of observed joint attention gestures was counted for each session. The CCX was administered at all time points, however the data from the current analysis come from the first time point (T0). All transcripts were verified and coded for reliability (20%) and administration of the CCX was evaluated for fidelity of implementation (20%).

Imitation. Motor imitation with objects was measured using the subscale of the Screening Tool for Autism in Young Toddlers and Children (STAT; Stone & Ousley, 1997). Four items were administered to all participants at screening to determine the child's ability to imitate actions with objects. A score of 0-4 was assigned based on the child's response.

Data analysis

The analyses for the first objective utilized a latent growth curve modeling approach within a structural equation-modeling framework (Bollen & Curran, 2006). This approach explains between-individual variability using time-invariant covariates. This approach also allows for identification of inter-individual differences in change with estimates of error.

Another benefit of this approach is that it allows for accommodation of missing data and specific tests of model fit. When the intercept it re-centered at a theoretically important time point (such as at the end of intervention), the intercepts can be interpreted as the average performance at the time point. Slope estimates allow us to interpret the change in a variable over time, or in a polynomial model, in the instantaneous rate of change where the intercept is centered. In a multi-group model, where separate growth curves are estimated for each group, the difference in intercept or slope estimates between groups can be tested for equality by allowing those estimates to be fixed or freed and then compared using the chi-square difference test. All computations, plotting and regression models were analyzed using R-studio running R version 3.2.3, and using mplus.R source code (R-Studio Team, 2015; R-Core Team, 2015; Fischer & Pau, 2015). All growth curve models were run using Mplus version 7 (Muthén, L. K., & Muthén, 1998).

Before specific examination of the research questions, the data were examined for outliers and non-normal distributions. For each variable proposed in the outlined analyses, a histogram created and was used to examine outliers. Outliers were verified for accuracy. Additionally, each variable set was examined for non-normality. All proposed baseline covariates were examined using a correlation matrix (Table 4). When variables were highly correlated, the variable with the strongest theoretical or empirical support from the literature for being included in the model was retained.

Prior to any modeling of the data, a spaghetti plot of the data was constructed for all participants (Figure 1 and Figure 3). By visually analyzing this plot, the functional form of the potential growth model could be better understood. Once the proposed model approach was identified, it was confirmed during the modeling building process using chi-square difference tests comparing simpler versus more complex functional models of growth. The most parsimonious model with good fit was retained. Good fit was determined by a non-significant chi-square test indicating that the null of good fit could not be rejected. Model fit was also evaluated by a Root Mean Square Estimate of Approximation (RMSEA; Steiger & Lind, 1980) confidence interval lower bound less than 0.05 and by observing a smaller information criteria (AIC; Akaike, 1987).

After the unconditional model was selected, the covariates were simultaneously added to the model and fixed across groups for intercepts and subsequently for linear slopes. Each covariate was allowed to be freely estimated across groups in order of theoretical relevance to the outcome (receptive language, speech sounds, play, joint attention, and imitation). Each pair of models with fixed and free covariate estimates across groups was compared using the chi-square difference test. A significant result indicates significant model fit improvement of the less restricted model, thus indicating an interaction with the covariate across groups.

Objective 1

To examine the interaction between the intervention group assignment and pre-intervention child characteristics, a conditional latent-growth curve model was used (Figure 1). Data from both language delay groups and the typical comparison sample, across six time-points, were used to fit the trajectory of growth of language diversity (NDW) and overall language (PLS) separately. A random intercept and random slopes model was fit to allow for outcome variability between and within participants. The appropriateness of across-group constraints on growth parameters was examined by comparing the fit indices of models with and without across-group constraints. Time-invariant covariates were added to the model to investigate whether they predicted inter-individual variability at T3 and T5 across and between the treatment and control groups. The time-invariant covariates considered in this model included receptive language, speech sound repertoire, play complexity, object motor imitation, and joint attention during the parent-child interaction observation. Covariates were allowed to vary one at a time in order of theoretical importance (speech sound repertoire, receptive language, joint attention, play complexity, and imitation). The time specific residuals were fixed across time initially and then allowed to vary over time and across groups as needed.

Objective 2

Objective 2 addressed the extent to which improvements in expressive vocabulary diversity (as measured by NDW) observed during the first four weeks of intervention predicted the response to treatment (greater total PLS scores) at the end of intervention, and the extent to which this effect interacted with group assignment. A multiple regression model was examined using group to predict immediate outcomes (Equation 1). The change in vocabulary diversity (NDW change between T0 and T1) was added to the model (Equation 2) and all relevant covariates were included to determine the extent to which early improvement in NDW predicted PLS scores at T3 or T5, and to determine if this improvement also interacted with group over time (Equation 3).

CHAPTER III

RESULTS

Reliability

All language samples and CCX assessments were transcribed and coded to measure NDW and joint attention gestures. All of the language sample and CCX transcripts were verified during the coding process for accuracy. From the 20% of the randomly selected CCX transcripts that were assessed for reliability, reliability was estimated at 95% (range=80-100%) across codes including gesture use. Fidelity of administration and accuracy in scoring the PLS was rated for 20% of the sample, distributed across time points. Correctness in administration and scoring averaged 95% (range=79-100%).

Speech sounds were rated by three speech-language pathologists who were previously trained in speech transcription and coding. All three speech-language pathologists reached 90% agreement on coding individual speech sounds before coding transcripts independently A randomly selected 25% of language samples were rated for reliability of speech sound coding. A third coder independently rated any sessions where reliability between the two primary coders was below 80% speech sound agreement (N=1), and consensus on coding each sound was reached by the team. Overall agreement on the on the point-by-point rating of individual speech sounds was moderate (82%, range=74-97%).

Objective 1

NDW. The spaghetti plots for number of different words across time were observed to have a piecewise functional form, with the knot centered on the post-intervention time point

(T03). This was confirmed through the model building process, which first examined the most parsimonious model, a single group linear latent-growth curve model, for which the null hypothesis of good fit could be rejected ($\chi^2(21) = 181.482, p = 0.00$). Additionally, the null hypothesis of good fit could be rejected for a quadratic functional form across groups ($\chi^2(17) = 120 - 756, p = 0.00$). A single group piece-wise model with a knot at T3 resulted in improved fit ($\chi^2(14) = 23.180, p = 0.05$), however ,this did not account for differing variances and residuals between groups. Thus, the final model for the first objective resulted in a multi-group piece-wise unconditional latent growth curve model with a knot point at T3 ($\chi^2(33) = 43.867, p = 0.10$, see Table 5).

Before examining the conditional model, the main effects as reported by Kaiser and Roberts (2015) were replicated within the growth curve modeling framework. A significant difference in intercepts at T3 was observed between the intervention and control groups by comparing models in which the intercepts were first fixed and then allowed to vary between groups. The chi-square difference test indicated significant improvement in model fit when NDW was allowed to vary between groups at T3 ($\chi^2(1) = 10.10, p = 0.001$). Additionally, significant improvement in model fit was observed when the linear slope between T0 and T3 was allowed to vary across groups ($\chi^2(1) = 14.12, p < 0.001$). These findings indicated the participants in the intervention used more words following intervention, and they were learning new words at a faster rate compared to the control group. These results, however, were not maintained 1-year following intervention. There were no significant intercept or slope differences between groups ($\chi^2(1) = 1.984, p = 0.159;\ \chi^2(1) = 1.870, p = 0.171$). Interestingly, slope differences between the intervention group and the typical comparison sample were not significantly different during intervention ($\chi^2(1) = 1.137, p = 0.286$), indicating that children in the intervention group were growing at the pace of typical development during the intervention period. Conversely, the control group participants grew significantly slower than typical development ($\chi^2(1) = 15.166, p = 0.00$) during the same 3 month period.

Question 1A. What pre-intervention child characteristics interact with group assignment to predict expressive vocabulary (a proximal measure; NDW) immediately following intervention (T3)? The five proposed time invariant covariates were simultaneously added to the piece-wise growth curve model and each covariate was fixed across groups (Figure 2). Receptive language ($\beta = 0.481, SE = 0.184, p = 0.009$) and number of different speech sounds ($\beta = 1.968, SE = 0.216, p = 0.000$) significantly predicted the number of different words (NDW) used at T3 across groups. However, joint attention, imitation and play complexity did not significantly predict post-intervention vocabulary diversity (NDW) . Allowing each covariate to vary across groups did not result in an improved model fit for any of the proposed covariates ($\chi^2 = 0.001 - 2.006\ (1), p > 0.05$).

The linear slopes from the first piece of the piece-wise model were also allowed to vary by the five theoretically important covariates. A similar pattern of results was observed. Improved model fit was not observed by allowing any of the covariates to be freely estimated across groups; however, speech sounds ($\beta = 0.288, p = 0.016$) and receptive language ($\beta = 0.481, p = 0.00$) were significant predictors of growth in NDW during the intervention period (T0-T3) across groups (Table 7).

Question 1B. What pre-intervention child characteristics interact with group assignment to predict expressive vocabulary (a proximal measure; NDW) 12 months following intervention (T5)? The main effects of the intervention on vocabulary diversity (NDW) 12 months following intervention were first examined in the unconditional growth curve model with the intercept re-centered at T5 (12 month follow up). Allowing the intercept to be freely estimated across groups did not significantly improve model fit; there were no significant mean differences in vocabulary diversity (NDW) between groups 12 months following intervention ($\chi^2(1) = 2.02, p = 0.155$). Additionally, the rate of change in NDW across the 12 months following intervention did not significantly differ between groups ($\chi^2(1) = 1.26, p = 0.262$).

The five proposed time invariant covariates were simultaneously added to the piece-wise growth curve model and each covariate was fixed across groups. Receptive language ($\beta = 0.481, p = 0.009$) and speech sounds ($\beta = 1.968, p = 0.000$) continued to be robust predictors of NDW 12 months following intervention. Speech sounds ($\beta = -0.135, p = 0.005$), play complexity ($\beta = 0.347, p = 0.037$), and joint attention skills ($\beta = 0.257, p = 0.002$) were significant predictors of growth during the 12 month follow up period.

PLS-4. The spaghetti plots for the PLS-4 standard scores were observed to have a non-linear polynomial functional form across the four measurement points (T0, T3, T4 and T5; see Figure 3). This was confirmed through the model building process. The most parsimonious model, a multiple group linear latent-growth curve model was examined first; the null hypothesis of good fit for this model could be rejected ($\chi^2(13) = 77.774, p = 0.000$). Additionally, the null hypothesis of good fit could be rejected for a quadratic functional form in the intervention group alone ($\chi^2(6) = 16.140, p = 0.013$). A multiple group quadratic functional form also did not result in improved fit ($\chi^2(5) = 12.844, p = 0.025$). Thus, the final model for the first objective resulted in a multi-group cubic unconditional latent growth curve model ($\chi^2(4) = 7.418, p = 0.115$), see Table 10).

Before examining the conditional model for the PLS, the main effects of the intervention on a standardized measure of language (PLS) at T3 were examined. A significant difference in intercepts at T3 was observed between the intervention and control groups by comparing models in which the intercepts were first fixed and then allowed to vary between groups. The chi-square difference test indicated significant improvement in model fit when PLS intercepts were allowed to vary between groups at T3 ($\chi^2(1) = 4.31\ p = 0.038$). Additionally, significant improvement in model fit was observed when the instantaneous linear slope at T3 was allowed to vary across groups ($\chi^2(1) = 4.51, p < 0.033$). These results indicated that participants in the intervention improved their performance on the PLS from T0 to T3 as compared to the control group. However, this effect did not maintain 12 months following intervention for slope ($\chi^2(1) = 0.524, p = 0.469$) or instantaneous linear growth (χ^2(1)=0.498, p=0.480). Although the intervention and typical groups performed differently following intervention on the PLS, the intervention group's linear rate of change at T3 was significantly greater than the typical group's rate of change ($\chi^2(1)$ =11.742, p=0.001). Unlike the control group, the rate of change for the treatment group was not significantly different in rate of in the typical group at T3 ($\chi^2(1) = 0.476, p = 0.490$).

Question 2A. What pre-intervention child characteristics interact with group assignment to predict total language ability (a distal outcome; PLS-4) immediately following intervention (T3)? The five proposed covariates (receptive language, diversity of speech sounds, play complexity, imitation, and joint attention) were simultaneously added to the model and fixed across groups. Receptive language was a significant predictor of performance on the PLS when the intercept was centered immediately following intervention (T3). Speech sounds, play complexity, imitation ability, and joint attention skills did not significantly predict post-intervention performance on the PLS. Joint attention skills, however, were allowed to vary across groups due to the improved model fit, indicating that joint attention skills may predict post-intervention PLS scores differentially across groups ($\chi^2 = 3.329(1), p = 0.068$). When the covariates were used to predict instantaneous linear growth at T3, joint attention skills were the only significant predictor of growth ($\beta = 0.416, p = 0.006$), with acceleration held constant across groups (quadratic and cubic terms were fixed). None of the covariates improved model fit when allowed to vary across groups, implying there were no moderated treatment effects by group (Table 12).

Question 2B. What pre-intervention child characteristics interact with group assignment to predict growth in total language ability (distal measure; PLS-4) 12 months following intervention (T5)? The cubic multiple-group growth curve with time centered at T5, 12 months following intervention, demonstrated a similar pattern of results for standardized language performance (PLS) as for vocabulary diversity (NDW). The model in which the intervention and control group values at T5 were estimated freely did not result in any improved fit; the mean estimates were fixed across groups indicating the groups did not perform differently on a standardized measure of language (PLS) 12 months after intervention ($\chi^2(1) = 0.524, p = 0.469$). Additionally, there was no significant difference in the instantaneous rate of change 12 months following intervention ($\chi^2(1) = 0.498, p = 0.480$). Each of the covariates was similarly related to T5 intercept and linear slope outcomes as at T3 (Table 11 & Table 12).

Objective 2

Does improvement in expressive vocabulary (NDW) during the first 4-weeks (as measured by the change between T0 and T1) interact with group assignment to predict the total standardized language outcomes (PLS-4) immediately following intervention (T3 and T5). In the first objective, receptive language was the only significant intercept predictor of later PLS scores. Therefore, receptive language was retained as a covariate in the multiple regression for this objective. Growth in language use (NDW) during the first month of the study did not significantly predict later overall language ability (PLS) at T3 ($\beta = 0.142, p = 0.215$) or at T5 ($\beta = 0.161, p = 0.252$). This effect did not interact with group assignment, and the majority of the variance at T3 and T5 in this model was due to receptive language ability ($\beta = 1.396, p = 0.000$; $\beta = 1.746, p = 0.000$). It should be noted that 12 participants were excluded from the T3 analysis and 18 from the T5 analysis due to missing PLS scores during the follow-period[1].

[1] Due to the list-wise deletion of all participants with missing data on any of the included variables, this analysis was re-run in the LGM framework to allow for retention of all participants. This was not part of the original proposal, but since the analysis produced a different pattern of results, it is included in Appendix C.

CHAPTER IV

DISCUSSION

The primary results of a randomized control trial, which examined the effects of Enhanced Milieu Teaching and parent training on observed vocabulary diversity (NDW) and overall language ability (PLS) in toddlers with primary language delays, indicated differences between groups on both measures of language growth at the end of the intervention period. Although the current analysis did not identify any variables that moderated the treatment effect, identification of the predictors of growth across the treatment and intervention groups serves an important role in characterizing the range of skills in this population of children with early language delays that are associated with varied language outcomes. Separate analyses across language constructs demonstrated differential and robust predictors of language growth for expressive vocabulary and overall language. The predictors of growth and change in vocabulary diversity did not explain differential overall language outcomes across groups, but did indicate specific child characteristics that are likely to predict improvements in language acquisition for this population.

Vocabulary diversity

None of the proposed predictors moderated the effects of the treatment over time for vocabulary. Diversity of speech sounds and receptive language at baseline were related to better growth in vocabulary diversity over the first three months (corresponding to the period of intervention). Additionally, receptive language predicted continued growth over time during the year following intervention. After the intervention period, sustained growth in vocabulary diversity was related to play complexity and joint attention skill at baseline. These two skills may indicate better engagement with objects and people, which may facilitate long-term language learning. In sum, growth in diversity of vocabulary over time is influenced by children's receptive language skills and their speech-sound repertoire, while joint attention and play skills appear to influence overall language skills over time.

Overall language ability

The measure of overall language ability in this study, the PLS, estimates the child's expressive and receptive language abilities over time. Although the PLS is a standardized assessment and it is relatively difficult to demonstrate large gains on standardized assessments in intervention studies, the cubic growth curve model indicated a significant improvement in the PLS in the intervention group as compared to the control group over the 3 months of the intervention period in the study. When the language growth of the intervention group was compared to the growth observed in the sample of typically developing children, it was apparent that intervention group continued to grow more slowly than their typical peers.

The main effects of the intervention on the improved language ability across groups over time were primarily driven by children's receptive language ability at baseline. The rate of language growth during the intervention period (the first three months) was driven by joint attention abilities. Having stronger joint attention skills at baseline may have had a differential effect on children in the intervention and control groups. This is indicated by substantively different estimates across groups. Although this was not a statistically significant difference ($p = .07$), it may be a meaningful difference between groups. Children in the control group had greater language ability if they initially demonstrated more joint attention, while language ability for children in the intervention group was not predicted by their baseline joint attention ability.

Although this effect falls short of statistical significance in this study, the relationship between joint attention and language ability should be considered in future studies.

Covariates

Speech sound diversity. Children with early language delays who have a great number of different speech sounds made greater gains in vocabulary diversity in both the treatment and control groups. Although this result is intuitive, it is important to consider these findings in the context of the additional predictors. The change in vocabulary improvement was primarily explained by diversity in speech sounds, indicating the importance of this skill for this population of toddlers with language delays. Speech sounds diversity at baseline, however, did not influence growth in overall language abilities. The difference in findings for vocabulary and overall language is likely related to differences in underlying constructs of these two measures of language. Vocabulary diversity was an observational measure of expressive language use. A large speech production repertoire is needed for vocabulary diversity to improve over time. In contrast, the measure of overall language ability (the PLS) includes items assessing auditory comprehension or receptive language skills and elicited productive language. Changes in overall language ability may be influenced by speech skills, but given comprehensive scope of the standardized measure, this influence is relatively less. In light of the importance of early vocabulary development for later social and academic performance, it may be important to add intervention components to target improving speech sounds for children with primary language delays.

Receptive language. Vocabulary growth was driven by receptive language ability in both the treatment and control group. Although receptive language did not interact with the intervention, it is a malleable factor that could be addressed in future intervention studies by tailoring the intervention specifically to support the child's receptive language abilities. It is important to consider that children with stronger receptive language and greater diversity of speech sounds in both the treatment and control group made substantial improvements in vocabulary, but neither the intervention nor the control participants caught up with the typical development. Therefore, participants with strong receptive language and speech abilities may still require early intervention to ensure the rate of growth in vocabulary that is needed to catch up to typically developing peers.

Play Complexity. Play complexity did not predict immediate post-intervention language improvements, yet it did predict vocabulary growth 12 months after the intervention period for children in both the intervention and control groups. Play complexity may be associated with high levels of object and person engagement over time. Children who spontaneously engage in play with objects may present more opportunities for natural language support processes (adult modeling vocabulary, sustained turn taking). Alternatively, consistent with prior findings, children with higher play complexity may have greater symbolic representation skills that support language learning in a naturalistic language intervention (Yoder & Stone, 2006). Children with autism also benefit linguistically from specific play teaching (Kasari et al. 2006). Thus, future research should consider supporting the development of symbolic play skills specifically in this language-delayed population.

Joint Attention. Children who demonstrated a higher rate of joint attention gestures (point, show, or give) did not show greater vocabulary growth during the three-month intervention period in either the treatment or the control group. However, children who entered the study with stronger joint attention skills were more likely to show a higher rate of growth in overall language ability over 12 months following the intervention period (T3-T5). Joint attention was a robust predictor across groups, and this skill may relate to the play complexity outcomes to enhance a child's environmental language learning ability.

Additionally, children in the control group made greater improvements in overall language ability after the first three months of the study (T0-T3) if they demonstrated greater joint attention skills at baseline, as compared to children in the intervention group, who made improvements during T0-T3 regardless of their joint attention at baseline. Although this result does not reach significance, the pattern of results should be considered. It is possible that children in the intervention group made improvements regardless of their baseline joint attention skills because their parents were taught to systematically respond to all communication attempts. Having established joint attention behaviors was not as important when parents were attending to and reinforcing all specific and non-specific communicative attempts. Conversely, knowing that teaching and supporting joint attention behaviors systematically allows for greater language learning over time (Kasari et al., 2006), it is possible that children in the intervention group received more specific supports for joint attention. However, children in the control group without strong joint attention behaviors may have received fewer or less well-matched responses from caregivers, which may have slowed their rate of language growth over time. Future research should consider how joint attention might be an important moderating variable in child language outcomes.

Imitation. Children with language delays who exhibited more imitation did not show greater growth in vocabulary diversity or language skills over time than children with who demonstrated less imitation. Consistent with recent findings of predictors of language growth in children with autism, motor imitation ability may not drive language learning over time (Yoder, Watson, & Lambert, 2015). However, the way in which imitation was measured (only four trials in the assessment of imitation) may have limited the variability in this score and thus, constrained generalizations about the relationship between imitation and language growth.

Response to treatment.

The rate at which children with language delays learn to use a diverse vocabulary (NDW) may not be associated with overall improvement in language ability (PLS) over time, and this effect does not interact with group assignment. These results may indicate that an improvement in vocabulary diversity may not be the best measure of early response to treatment in this population. Because NDW may estimate a single underlying construct of language (expressive vocabulary), a more comprehensive language measure may be necessary to establish an early read of response to treatment intended to impact overall language development. This poses potential methodological limitations. Standardized measures may not be sensitive to relatively small changes in language skills since items represent a sample of skills at a particular age rather than a continuous measure of skill development. Further, standardized tests cannot be administered more frequently than recommended by their developers without compromising the validity of the measure. In adaptive treatment studies, two alternative approaches have been used to assess early response to treatment : assessing progress on a range of measures representing different aspects of language and evaluating response to treatment based on meeting criterion levels of change across the majority of the measures (Kasari, et al 2014) or clinician and research team ratings of clinical global impressions of progress in treatment (Kasari et al, in progress). This is an important area for further research examining alternative measures of response to treatment that predict later outcomes.

Alternative predictors.

While the current analysis did not indicate any moderated treatment effects, it is possible that the differential outcomes across and within groups were driven by factors other than child characteristics, particularly parent characteristics (e.g., responsiveness, linguistic input, or education) or environmental variables (e.g., socio-economic status, home literacy environment) that were not measured in this study. Although additional analyses are needed to explore the contributions of parent and environment variables on child language outcomes, this finding that none of the theoretically driven child characteristics moderated treatment effects may be important for this population. Adult behavior supporting language or environmental factors may be more malleable than some child characteristics.

Limitations.

The results of this study must be considered in perspective of the study's limitations. First, the sample size, although larger than many other studies with children with primary language delays, is still relatively small. It may be that the small but meaningful impact of some theoretically important covariates could not be observed in this study do to the sample size. Second, similar to all predictive studies, this study cannot draw precise conclusions about direct relationships between variables. Without the specific manipulation of each covariate, it remains plausible that an unidentified confounding variable may be driving the observed relationships in this study. Third, more accurate measures of key behaviors could improve specificity in estimating the key behaviors of each covariate. For example, a more systematic probe of joint attention, such as the Early Social Communication Scale (Mundy, Delgado, Block, Venezia, Hogan, & Seibert, 2003), ability could provide a more representative estimate of children's joint attention ability. Additionally, more opportunities to demonstrate motor imitation could also allow for a more stable measure of imitation ability. Although ASD symptomatology was measured in this study, the large amount of missing data prevented including this as a covariate. Future research with this population should include a sensitive measure of ASD to estimate the effect of ASD symptomatology on language growth over time.

Future Research.

The results of this study suggest a need for additional research in this area. Primarily, this study should be replicated with a larger sample to better estimate the main effects of the intervention and to allow for a better estimate of the relationship between child, parent, and environmental variables and child language growth within and across treatment arms. A larger study should also incorporate more precise estimates of joint attention and imitation as well as a measure of ASD symptomatology. Additional theoretically important measures should also be considered in future research such as non-verbal cognitive ability, response to joint attention, and additional measures of play and engagement.

Another goal for future research is to develop an adaptive intervention schema to specifically incorporate key predictors of language growth in children with primary language delays. Perhaps teaching joint attention, receptive language, and speech sounds prior to a naturalistic intervention would result in greater accelerated language learning. The specific dose and density of strategies necessary to gain the desired improvements in children with primary language delays could be investigated using an adaptive treatment study design including key variables such as receptive language, speech sounds, and joint attention as tailoring variables .

Investigators also should examine the specific relationship between joint attention skills and language during adult-child interactions. A sequential analysis could help explain the relationship between joint attention and adult language support strategies. The relationship between joint attention and intervention strategies would be important to understand as well as the relationship between parent language support strategies and joint attention outside of the context of the intervention.

Recommendations for Practice

Although substantial gains were observed in overall language ability and natural use of language in the intervention group as compared to the control group, the trajectory of growth was not sustained in the treatment group after the end intervention. The accelerated rate of growth observed during the three months treatment would need to be maintained over a longer course of time to facilitate sufficient gains in language to catch up to children with typical development. Additionally, intervention components may need to be added or adapted to better facilitate response to treatment for all children. Although the following recommendations should be established empirically, the results of this study provide enough evidence to support implications for early intervention.

First, a longer intervention period is likely to aid in persistent language growth. The intervention in the current study was extremely brief (3 months) and resulted in significant gains in language. Alternatively, the rate of language growth might be maintained following treatment with use of intermittent booster sessions rather than continuous intervention. If the intervention effects are truly driven by parent or environmental factors, maintenance sessions that include parent training, teaching strategies to address advanced vocabulary, morphology and syntax targets may support continuity in intervention effects across time.

Second, early interventionists should utilize additional intervention strategies to support receptive language, joint attention, and speech sound diversity in children with primary language delays. These skills play an important role in immediate and long-term outcomes of language and vocabulary development and should be facilitated. Although future researchers will need to establish empirically based adaptations to Enhanced Milieu Teaching, it is clear that in the interim, early interventionists could better facilitate growth in these additional areas. Teaching communicative partners to be responsive regardless of the quality of the child's initiations may also be an important strategy to support communication and language growth in this population.

Conclusions

Young children with primary language delays demonstrated significant improvement in overall language ability as well as observed vocabulary use immediately following an therapist and parent implemented intervention. None of the child characteristics evaluated in this study moderated the treatment outcome. Across the treatment and control groups, children who had better receptive language abilities, joint attention skills, and speech sound diversity exhibited better growth in language over time, but all children with language delays continued to grow more slowly than typically developing children. While children with low receptive language, fewer joint attention behaviors and less diverse speech sounds may be at relatively greater risk for poor language outcomes, the results of the study suggest that the population of children with delayed language remain at risk for delayed development . For a population for whom the recommendation is, "wait and see" if a child improves, this study indicates that waiting will not result in sufficient improvements in language abilities such that these children catch up to their typically developing peers. There is a continuing need to development interventions that target the predictors of language development in this population as potential drivers of language development and to test sustained interventions that support a range of language skills in the toddler and preschool years.

EQUATIONS

Equation 1

$$PLS_{t3} = b_0 + b_1 Group + b_2 NDW_{T1-T0} + covariates + e$$

Equation 2

$$PLS_{t3} = b_0 + b_1 Group + b_2 NDW_{T1-T0} + b_3 Group * NDW_{T1-T0} + covariates + e$$

Table 1 Prior studies with children who have primary language delays

Study	Design	Intervention	Population	Effects	Notes
Law et al. 1999	RCT	Low-dose parent training	Toddlers with expressive and receptive delays	No effect on language outcomes	Strong effects on parent report measures of vocabulary
Glogowska et al., 2000	Waitlist	Low-dose, addition of eclectic SLP services	Preschoolers with expressive and receptive delays	No effect on spoken language or phonology	Not much evidence for spontaneous recovery
Buschmann, 2009	RCT	Parent implemented 3 month intervention: Shared book reading, modeling, and prompting	Toddlers with expressive delay only (61, but 47 in analysis)	Significant differences in language	26% of control remained delayed compared to 8% intervention
Wake et al., 2011	RCT	Low-dose parent training which included 6 group sessions	Toddlers who were in 20^{th} percentile or lower for language at 18 months	No language effects	Very low dose and no measure of fidelity
Roberts & Kaiser, 2015	RCT	clinician + parent naturalistic intervention	97 toddlers, 30 months old, language delays and typical cognition	Significant improvement on proximal measures (NDW) not on distal (PLS)	Some children did make substantial gains– who are these children?

Table 2 Participant descriptions

	Intervention	Control*	Typical sample
n	45	52	88
Age, months	30 (5.2)	30 (4.9)	29.2 (5.8)
Cognitive, Bayley Composite	91 (8.4)	88 (7.4)	104.15 (9.48)
Bayley Language Comprehension	75.93 (8.60)	74.88 (7.88)	113.65 (14.14)
Percent Male	82%	81%	82%
Percent Minority	22%	21%	25%
PLS-receptive	76.49 (17.35)	73.12 (14.82)	89.87 (23.61)
PLS-expressive	75 (7.44)	74.98 (6.91)	91.41 (25.24)
NDW, T00	19.04 (17.94)	16.54 (16.81)	49.92 (45.63)
Imitation: STAT [0-4]	3 (1.18)	2.67 (1.34)	3.75 (0.26)
ADOS social subscale	2.78 (3.89)	4.04 (4.48)	
Average highest level of play (SPA) [1-16]	7.93 (2.05)	8.36 (1.85)	8.70 (1.87)
Joint attention during CCX	6.66 (3.86)	6.34 (4.20)	

*two participants removed from all following analyses due to missing data across all measures except IQ

Note. No significant differences were observed between Intervention and Community groups during the T0 assessment period

Table 3 Constructs by Measure over time

Construct	Assessment	T0: 0-months	T1: 1-month	T2: 2-months	T3: 3-months	T4: 9- months	T5: 15-months
NDW	Language sample (Roberts & Kaiser, 2015)	X	X	X	X	X	X
Standardized language	PLS-4 (Zimmerman et al., 2011)	X			X	X	X
IQ	Bayley (Bayley, 1993)	X					
Receptive language	Bayley-receptive communication (Bayley, 1993)	X X					
Speech sound repertoire	Language sample (Roberts & Kaiser, 2015)	X					
Imitation	STAT (Stone & Ousley, 1997)	 X					
Joint Attention	Caregiver-child interaction (Roberts & Kaiser, 2015)	X					
Complexity of Play	SPA (Ungerer & Sigman, 1981)	X					

Table 4 Correlation Matrix of Pre-Intervention Characteristics

	NDW	PLS	PLS-AC	Bayley Receptive	IQ	Speech	Play	Imitation	CBCL	Complexity	Age	PPVT	EOW	MCDI	Joint Attention
NDW	1.00														
PLS	0.13	1.00													
PLS-AC	-0.04	0.92	1.00												
Bayley Receptive	0.25	0.77	0.78	1.00											
IQ	-0.03	0.68	0.71	0.57	1.00										
Speech	0.71	0.03	-0.14	0.09	-0.12	1.00									
Play	0.25	0.37	0.33	0.40	0.34	0.20	1.00								
Imitation	0.13	0.23	0.25	0.14	0.22	0.05	0.29	1.00							
CBCL	0.15	-0.12	-0.21	-0.04	-0.08	0.22	0.20	-0.05	1.00						
Complexity	0.55	-0.01	-0.17	0.00	-0.13	0.87	0.17	0.07	0.28	1.00					
Age	0.55	-0.31	-0.31	-0.08	-0.28	0.43	0.30	0.12	0.16	0.35	1.00				
PPVT	0.15	0.70	0.71	0.57	0.51	0.10	-0.07	0.16	-0.14	0.08	-0.19	1.00			
EOW	0.59	0.28	0.11	0.28	0.04	0.38	0.63	0.18	0.18	0.34	0.28	0.18	1.00		
MCDI	0.74	0.01	-0.15	0.14	-0.15	0.47	0.19	0.11	0.27	0.40	0.60	-0.07	0.63	1.00	
Joint Attention	-0.03	0.18	0.22	0.24	0.14	-0.08	0.10	0.02	-0.01	-0.04	-0.05	0.15	0.00	-0.02	1.00

Table 5 NDW model building

Model	χ^2(df)	RMSEA	AIC
Single Group, linear LGM	181.482 (21), p=0.000	0.284, p=0.000	4685.282
Single group, quadratic LGM	120.756(17), p=0.000	0.253, p=0.000	4632.556
Single group piece-wise (knot=T3) LGM	23.180(14), p=0.050	0.080, p=0.187	4540.980
Multiple Group, piece-wise (knot=T3) LGM	43.867(33), p=0.100	0.083, p=0.210	4538.562

Table 6 Predictors of NDW intercept at T3

Covariate	estimate	SE	p	X^2 difference (group interaction)
Receptive	0.481	0.184	0.009	0.039(1), p=0.843
Speech	1.968	0.216	0.000	0.001(1), p=0.974
Play	0.233	0.766	0.761	0.442(1), p=0.151
Joint Attention	-0.474	0.396	0.231	0.442(1), p=0.506
Imitation	1.078	1.034	0.297	0.056(1), p=0.813

Table 7 Predictors of NDW linear slope between baseline and T3

Covariate	estimate	SE	p	X^2 difference (group interaction)
Receptive	0.481	0.184	0.000	0.045(1), p=0.832
Speech	0.288	0.120	0.016	0.812(1), p=0.368
Play	-0.157	0.426	0.463	0(1), p=1.000
Joint Attention	-1.148	0.773	0.138	0.262(1), p=0.609
Imitation	2.226	2.167	0.304	0.245(1), 0.621

Table 8 Predictors of NDW Intercept differences at T5

Covariate	estimate	SE	p	X^2 difference (group interaction)
Receptive	0.481	0.184	0.009	0.039(1), p=0.843
Speech	1.968	0.216	0.000	0.001(1), p=0.974
Play	0.232	0.766	0.762	2.066(1), p=0.151
Joint Attention	-0.474	0.396	0.231	0.442(1), p=0.506
Imitation	1.078	1.034	0.297	0.056(1),p=0.813

Table 9 Predictors of NDW growth between T3 and T5

Covariate	estimate	SE	p	X^2 difference (group interaction)
Receptive	-0.044	0.043	0.301	2.263(1), p=0.132
Speech	-0.135	0.049	0.005	1.707(1), p=0.191
Play	0.347	0.166	0.037	1.845(1), p=0.174
Joint attention	0.257	0.084	0.002	0.115(1), p=0.735
Imitation	-0.089	0.084	0.716	0.635(1), p=0.426

Table 10 Model building PLS

Model	χ^2(df)	RMSEA	AIC
Multiple Group, linear LGM	77.774 (13), p=0.000	0.324, p=0.000	2638.653
Multiple Group, quadratic intervention LGM	16.140(6), p=0.013	0.189, p=0.025	2591.019
Multiple Group, quadratic LGM	12.844(5), p=0.025	0.182, p=0.042	2589.72
Multiple Group, cubic LGM	7.418(4), p=0.115	0.134, p=0.156	2526.89

Table 11 Predictors of PLS intercept at T3

Covariate	estimate	SE	p	X^2 difference (group interaction)
Receptive	0.921	0.103	0.000	0.306(1), p=0.580
Speech	-0.084	0.115	0.465	0.073(1), p=0.787
Play	0.495	0.418	0.237	0.812(1), p=0.368
Imitation	0.814	0.576	0.157	0.725(1), p=0.395
Joint Attention	0.044	0.224	0.195	
Intervention	-0.424	0.327	0.194	3.329(1), p=0.068
Control	0.362	0.275	0.189	

Note. Allowing covariates to be freely estimated across groups did not result in improved model fit, as observed in the χ^2difference test, however between group estimates for joint attention are reported to allow for substantive interpretation

Table 12 Predictors of linear slope for PLS scores at T3

Covariate	estimate	SE	p	X^2 difference (group interaction)
Receptive	0.130	0.077	0.091	0.892(1), p=0.345
Speech	-0.131	0.085	0.122	0.920(1), p=0.337
Play	0.305	0.307	0.320	1.665(1), p=0.197
Imitation	0.683	0.434	0.116	0.977(1), p=0.323
Joint Attention	0.416	0.150	0.006	0.021(1), p=0.884

Table 13 Predictors of PLS intercept at T5

Covariate	estimate	SE	p	X^2 difference (group interaction)
Receptive	0.976	0.106	0.000	0.306(1), p=0.580
Speech	-0.080	0.112	0.474	0.073(1), p=0.787
Play	0.495	0.418	0.237	0.812(1), p=0.368
Imitation	0.814	0.576	0.157	0.725(1), p=0.395
Joint Attention	0.044	0.224	0.845	
Intervention	-0.424	0.327	0.194	3.329(1), p=0.068
Control	0.362	0.275	0.189	

Note. Allowing covariates to be freely estimated across groups did not result in improved model fit, as observed in the χ^2difference test, however between group estimates for joint attention are reported to allow for substantive interpretation

Table 14 Predictors of linear slope for PLS scores at T5

Covariate	estimate	SE	p	X^2 difference (group interaction)
Receptive	0.107	0.077	0.165	0.892(1), p=0.345
Speech	-0.125	0.083	0.131	0.920(1), p=0.337
Play	0.344	0.304	0.257	1.665(1), p=0.197
Imitation	0.717	0.430	0.095	0.977(1), p=0.323
Joint Attention	0.416	0.150	0.002	0.021(1), p=0.884

Table 15 Regression results for Predictors of T3 PLS scores

	Estimate	Standard Error	P
Intercept	-29.099	12.203	0.019*
Change T0->T1	0.124	0.099	0.215
Group: Intervention	5.026	2.649	0.061
Receptive	1.396	0.162	0.000*

(12 observations removed due to missingness)
Multiple R-squared: 0.5288
Adjusted R-squared: 0.5114
F-statistic: 30.3 on 3 and 81 DF, p-value: 0.00

Table 16 Regression Results for Group Interaction at T3

	Estimate	Standard Error	P
Intercept	-29.3835	12.342	0.020*
Change T0->T1	0.149	0.151	0.327
Group: Intervention	5.406	3.402	0.327
Receptive	1.396	0.163	0.000*
Change*Group	-0.044	0.197	0.825

(12 observations deleted due to missingness)
Multiple R-squared: 0.5291,
Adjusted R-squared: 0.506
F-statistic: 22.47 on 4 and 80 DF, p-value: 0.00

Table 17 Regression Results for Predictors of T5 PLS

	Estimate	Standard Error	P
Intercept	-44.975	16.899	0.009*
Change T0->T1	0.1608	0.1392	0.252
Group: Intervention	1.554	3.636	0.670
Receptive	1.746	0.222	0.000*

(12 observations deleted due to missingness)
Multiple R-squared: 0.4809,
Adjusted R-squared: 0.4601
F-statistic: 27.16 on 3 and 75 DF, p-value: 0.000

Table 18 Regression Results for Group Interaction at T5

	Estimate	Standard Error	P
Intercept	-43.843	17.045	0.012*
Change T0->T1	0.062	0.202	0.757
Group: Intervention	-0.494	4.761	0.917
Receptive	1.7434	0.224	0.000*
Change*Group	0.184	0.275	0.505

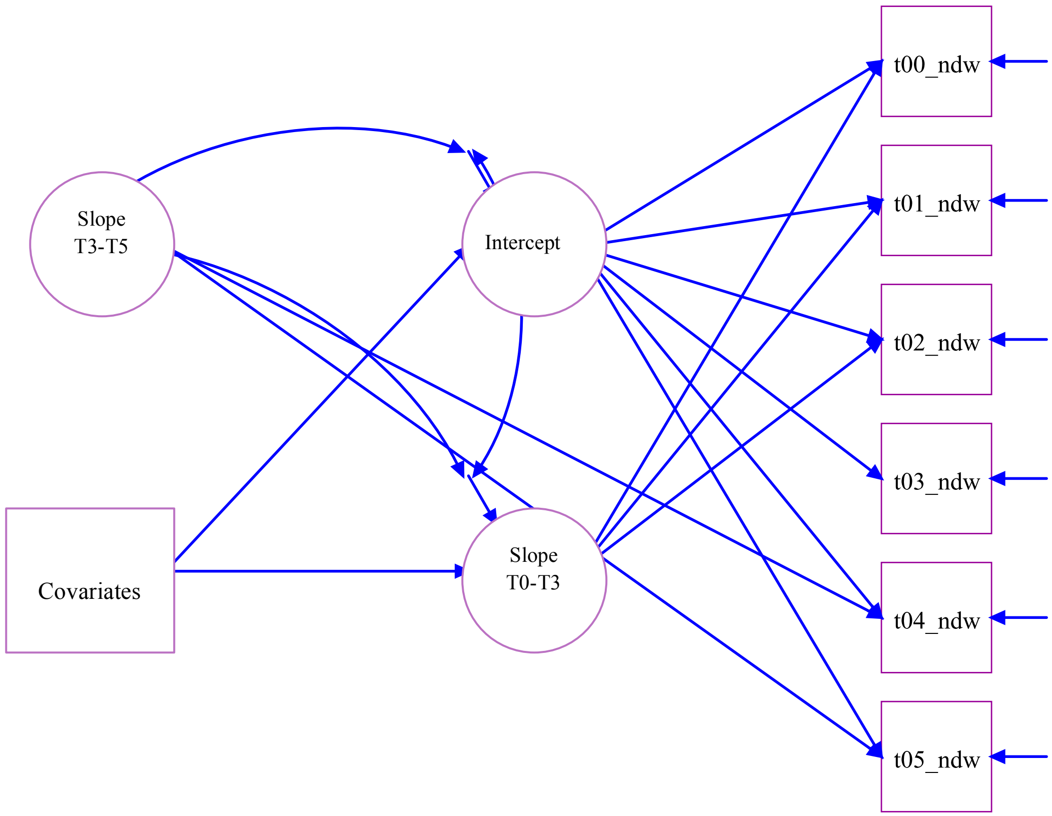

Figure 1. Path diagram of conditional model outlining first research question

**covariates merged to improve readability, however each covariate is assumed to have it's own error and influence on the slope and intercept*

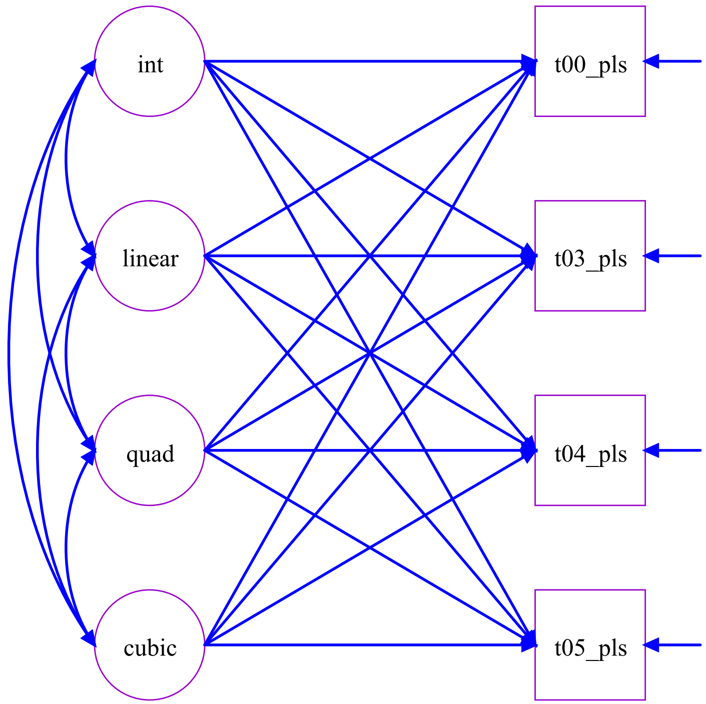

Figure 2. Path diagram of unconditional model outlining second research question

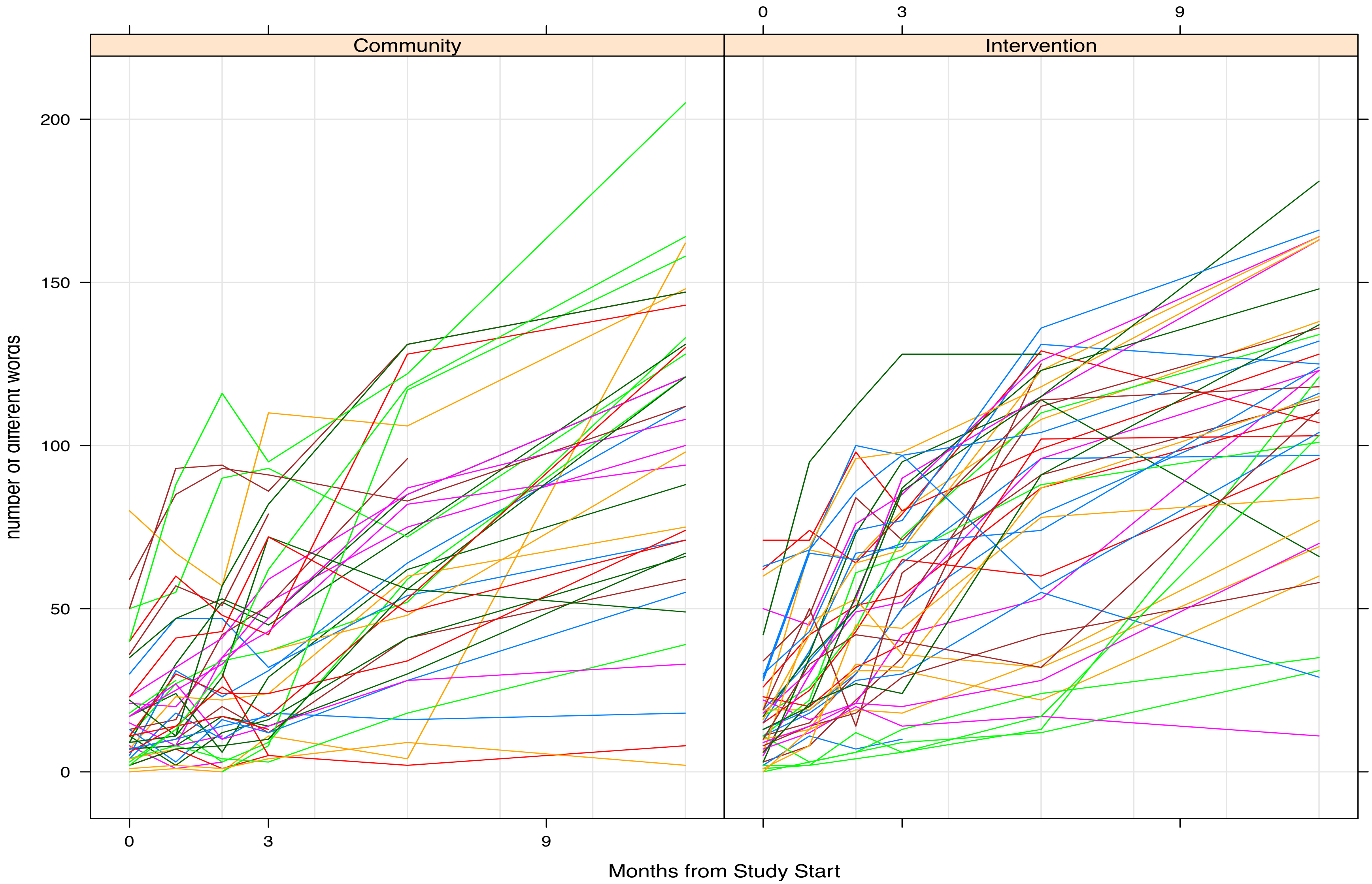

Figure 3. Spaghetti plot: Number of Different Words over time

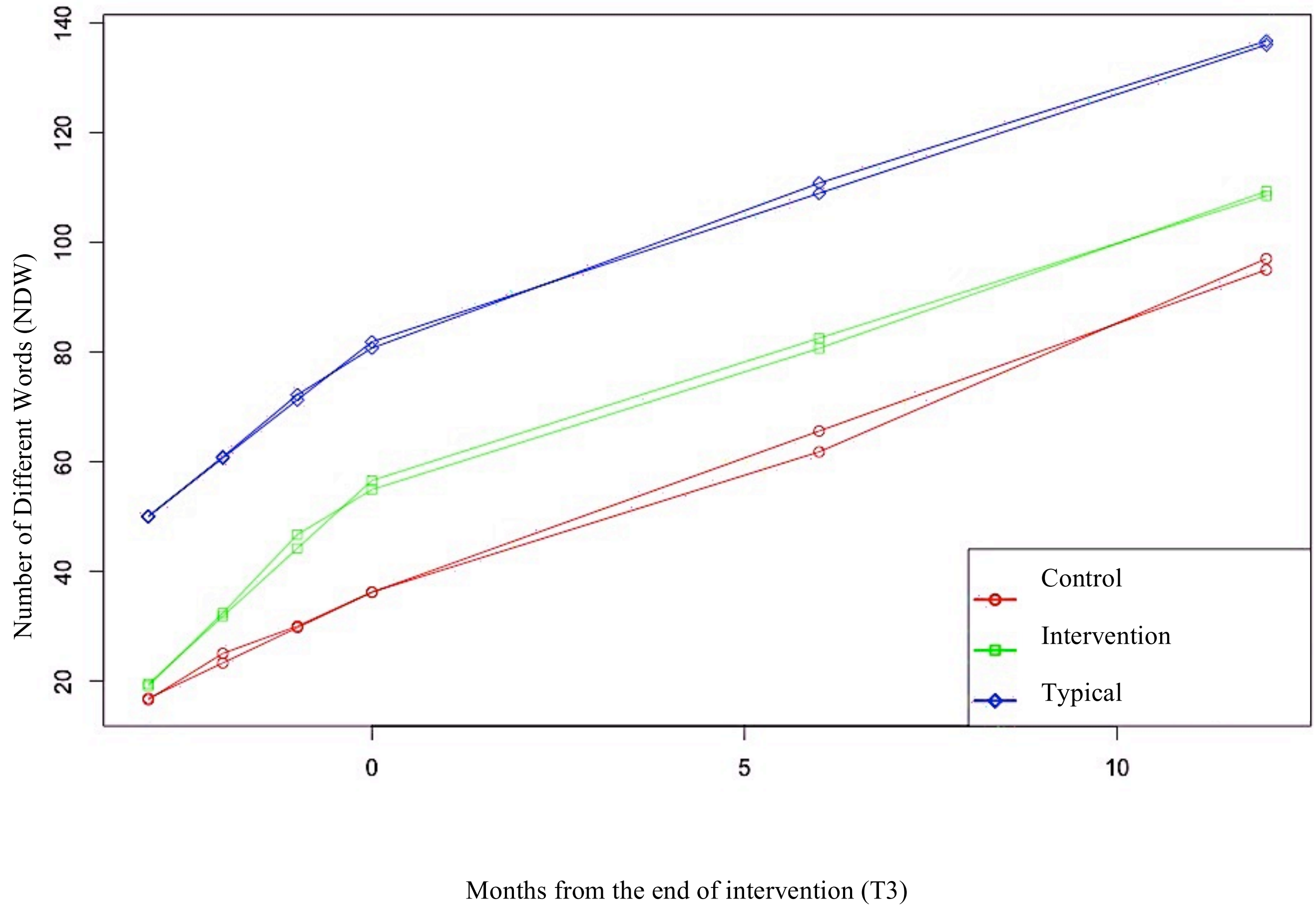

Figure 4 Observed and Estimated Means of NDW over time

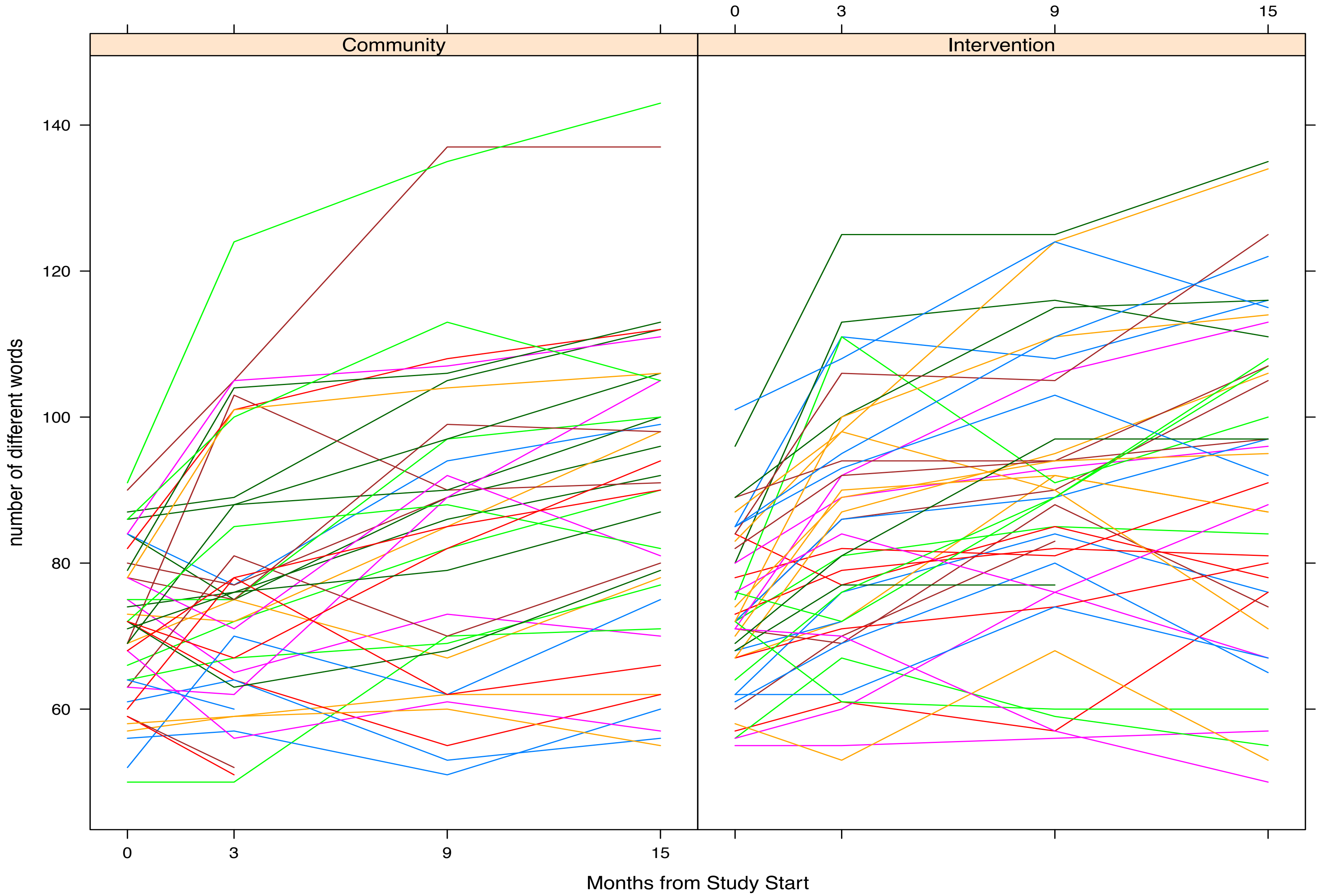

Figure 5. Spaghetti plot of observed trajectories on PLS

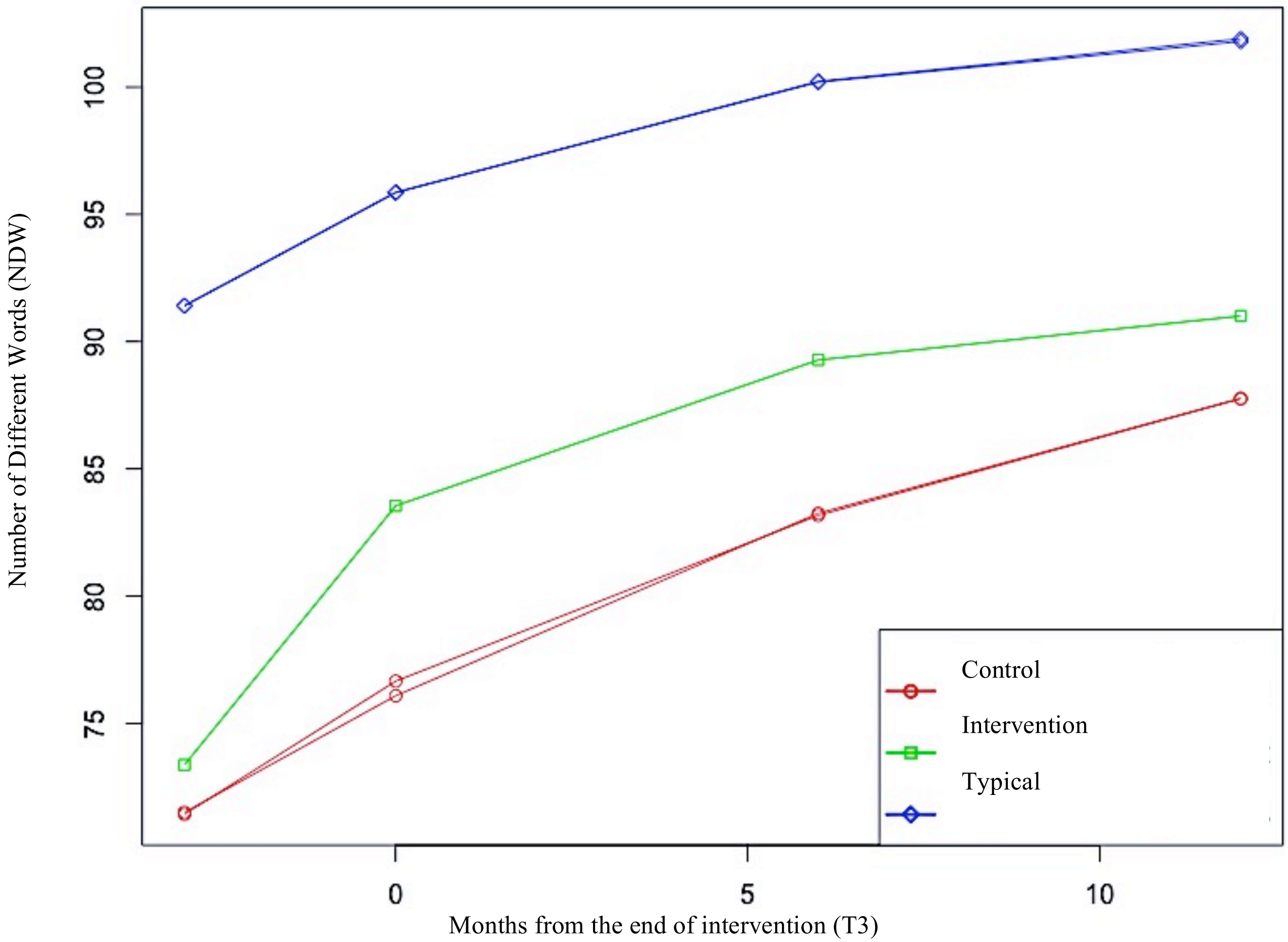

Figure 6 Observed and Estimated Means for PLS

Appendix A: Kidtalk Coding Manual: Independence only

Independence: the level of support the child needs to communicate.

1. **Unprompted [U]:** the child is spontaneously communicating. The child is communicating out of his/her own will and/or is initiating communication with the adult. If the child imitates part or all of the adult's preceding communication, but changes the form of communication OR adds content word(s), it is considered unprompted [u].

2. **Imitated [I]:** the child imitates all or part of the preceding adult utterance and does not add any content. If the child imitates all or part of the preceding adult utterance but adds content word(s), changes the mode of communication, or changes the function of the word/gesture, code [u]. If the adult asks a question and the child imitates part or all of the adult's utterance (does not add words), code [i].

3. **Prompted [P]:** the child is communicating in response to a Say or Choice prompt. The child must be saying the prompted words. If the child says something unrelated in response to a Say or Choice prompt, code [e].

4. **Supported/Elicited [E]:** the child is using spontaneous language in response to an adult communication prompt/cue. Child utterances in response to any of the following will receive this code:
   (a) Open Question
   (b) Yes/No Question
   (c) Clarifying question
   (d) Test Question
   (e) Time Delay*

   *Only code child responses to a Time Delay [e] if the child's communication is related to the Time Delay. If the child is spontaneously communicating about something unrelated during a Time Delay, code [u].

   NOTE: If the child imitates all or part of the adult's utterance (and does not add words), code [i].

NOTE:
Child unintelligible [cx] do not count as words added by the child that make the utterance unprompted [u]. If the only intelligible words in the child's utterance are imitated words from the adult's previous utterance, the child's utterance will be considered imitated [i].

Appendix B: Speech Sound Coding Sheet

ID: Coder: Date Coded:

Length of video:

Phonetic Inventory

Stops	1	+	Fricatives	1	+
/p/			/f/		
/b/			/v/		
/t/			/s/		
/d/			/z/		
/k/			/ʃ/		
/g/			/ʒ/		
Glides			/θ/		
/w/			/ð/		
/j/			/h/		
Liquids			**Affricates**		
/l/			/tʃ/		
/r/			/dʒ/		
Nasals			**Vowels**		
/m/			/i/		
/n/			/e/		
/ŋ/			/u/		
			/o/		

Total 1 occurance:

Total + 1 occurance

Syllable Structure

	1	+	Notes
Vowels			
CV			
CVC			
CVCV			
Variegated			
Reduplicated			
Multi-syllabic			

Silence

	Yes	No		Yes	No
0:00-0:29			5:00-5:29		
0:30-0:59			5:30-5:59		
1:00-1:29			6:00-6:29		
1:30-1:59			6:30-6:59		
2:00-2:29			7:00-7:29		
2:30-2:59			7:30-7:59		
3:00-3:29			8:00-8:29		
3:30-3:59			8:30-8:59		
4:00-4:29			9:00-9:29		
4:30-4:59			9:30-9:59		

Total intervals of silence:

Appendix C: Alternative Analysis of Objective 2

The second objective was also analyzed within the multiple group linear piece-wise model. This was done to retain more cases. The slope in NDW vocabulary growth during the course of the intervention across groups was estimated using the final language performance on the PLS as a covariate. This model estimates if the rate of growth associated with later PLS, and if this relationship varies across groups. Receptive language was retained as an intercept predictor, similar to the multiple regression analysis.

Unlike the multiple regression results, the multi-group conditional linear piece-wise model indicates that PLS scores 1-year following intervention are significantly related to the amount of vocabulary growth (NDW) during intervention ($\beta = 0.215, SE = 0.033, p = 0.000$). This relationship does not produce and interaction with group such that there is not a significantly better model fit if the slopes are allowed to vary across groups ($\chi^2(1) = 0.679, p = 0.410$). Therefore, although growth in vocabulary may be an important indicator of later language ability, this relationship does not moderate treatment outcomes.

REFERENCES

Akaike, H. (1987). Factor analysis and AIC. Psychometrika, 52(3), 317-332.

Almirall, D., Collins, L. M., & Murphy, S. A. (2012). Getting SMART about Developing Individualized Sequences of Health Interventions. In *ANNALS OF BEHAVIORAL MEDICINE* (Vol. 43, Pp. S59–S59). SPRINGER 233 SPRING ST, NEW YORK, NY 10013 USA.

Bayley, N. (1993). *Bayley scales of infant development: Manual*. Psychological Corporation.

Bishop, D. V. M., & Leonard, L. (2014). *Speech and Language Impairments in Children: Causes, Characteristics, Intervention and Outcome*. Psychology Press.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective* (Vol. 467). John Wiley & Sons.

Bopp, K. D., & Mirenda, P. (2011). Prelinguistic predictors of language development in children with autism spectrum disorders over four–five years. *Journal of Child Language*, *38*(3), 485.

Buschmann. (2009). Parent based language intervention for 2-year-old children with specific expressive language delay: a randomized controlled trial., *94*(2), 110–116.

Brady, N. C., Marquis, J., Fleming, K., & McLean, L. (2004). Prelinguistic predictors of language growth in children with developmental disabilities. *Journal of Speech, Language, and Hearing Research*, *47*(3), 663–677.

Buschmann, A., Jooss, B., Rupp, A., Dockter, S., Blaschtikowitz, H., Heggen, I., & Pietz, J. (2008). Children with Developmental Language Delay at 24 Months of Age: Results of a Diagnostic Work-Up. *Developmental Medicine & Child Neurology*, *50*(3), 223–229.

Chiat, S., & Roy, P. (2008). Early phonological and sociocognitive skills as predictors of later language and social communication outcomes. *Journal of Child Psychology and Psychiatry*, *49*(6), 635–645. http://doi.org/10.1111/j.1469-7610.2008.01881.x

Collins, L. M., Murphy, S. A., & Strecher, V. (2007). The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. *American Journal of Preventive Medicine*, *32*(5), S112–S118.

Deconinck, N., Soncarrieu, M., & Dan, B. (2013). Toward Better Recognition of Early Predictors for Autism Spectrum Disorders. *Pediatric Neurology*, *49*(4), 225–231.

Feldman, H. M., Dale, P. S., Campbell, T. F., Colborn, D. K., Kurs-Lasky, M., Rockette, H. E., & Paradise, J. L. (2005). Concurrent and Predictive Validity of Parent Reports of Child Language at Ages 2 and 3 Years. *Child Development*, *76*(4), 856–868. http://doi.org/10.1111/j.1467-8624.2005.00882.x

Bernd Fischer and Gregoire Pau (2015). rhdf5: HDF5 interface to R. R # package version 2.4.0.

Fricke, S., Bowyer□Crane, C., Haley, A. J., Hulme, C., & Snowling, M. J. (2013). Efficacy of language intervention in the early years. *Journal of Child Psychology and Psychiatry*, *54*(3), 280–290. http://doi.org/http://dx.doi.org/10.1111/jcpp.12010

Fudala, J. B., & Reynolds, W. M. (1986). *Arizona Articulation Proficiency Scale: Manual*. Western Psychological Services.

Fujiki, M., Brinton, B., Hart, C., & Fitzgerald, A. (1999). Peer Acceptance and Friendship in Children with Specific Lan... : Topics in Language Disorders. *Topics in Language Disorders*.

Fujiki, M., Brinton, B., Morgan, M., & Hart, C. H. (1999). Withdrawn and Sociable Behavior of Children With Language Impairment. *Language Speech and Hearing Services in Schools*, *30*(2), 183. http://doi.org/10.1044/0161-1461.3002.183

Glogowska, M., Roulstone, S., Enderby, P., & Peters, T. J. (2000). Randomised controlled trial of community based speech and language therapy in preschool children. *BMJ*, *321*(7266), 923.

Hawa, V. V., & Spanoudis, G. (2014). Toddlers with delayed expressive language: An overview of the characteristics, risk factors and language outcomes. *Research in Developmental Disabilities*, *35*(2), 400–407. http://doi.org/http://dx.doi.org.www2.lib.ku.edu/10.1016/j.ridd.2013.10.027

Joss, R. A., Alberto, P., Bleher, E. A., Ludwig, C., Siegenthaler, P., Martinelli, G., … Senn, H. J. (1994). Combined-modality treatment of small-cell lung cancer: randomized comparison of three induction chemotherapies followed by maintenance chemotherapy with or without radiotherapy to the chest. Swiss Group for Clinical Cancer Research (SAKK). *Annals of Oncology: Official Journal of the European Society for Medical Oncology / ESMO*, *5*(10), 921–928.

Kaiser, A. P., & Roberts, M. Y. (2013a). Parent-Implemented Enhanced Milieu Teaching With Preschool Children Who Have Intellectual Disabilities. *Journal of Speech, Language and Hearing Research*, *56*(1), 295.

Kaiser, A. P., & Roberts, M. Y. (2013b). Parents as Communication Partners: An Evidence-Based Strategy for Improving Parent Support for Language and Communication in Everyday Settings. Perspectives on Language Learning and Education, 20(3), 96–111.

Kasari, C., Freeman, S., & Paparella, T. (2006). Joint attention and symbolic play in young children with autism: A randomized controlled intervention study. Journal of Child Psychology and Psychiatry, 47(6), 611-620.

Kasari, C., Kaiser, A., Goods, K., Nietfeld, J., Mathy, P., Landa, R., ... & Almirall, D. (2014). Communication interventions for minimally verbal children with autism: A sequential multiple assignment randomized trial. Journal of the American Academy of Child & Adolescent Psychiatry, 53(6), 635-646.

Kasari, C., Kaiser, A.P., Lord, C., & Smith, T., (in progress). Adaptive Interventions for Minimally Verbal Children with ASD in the Community. National Institutes of Health: Autism Center of Excellence. Grant # NIH/NICHD RO1HD073975

Law, J., Kot, A., & Barnett, G. (1999). A comparison of two methods for providing intervention to three year old children with expressive/receptive language impairment. City University.

Leonard, L. B. (2014). Children with Specific Language Impairment. MIT Press.

Leyfer, O. T., Tager-Flusberg, H., Dowd, M., Tomblin, J. B., & Folstein, S. E. (2008). Overlap between autism and specific language impairment: Comparison of autism diagnostic interview and autism diagnostic observation schedule scores. Autism Research, 1(5), 284–296.

Lord, C., Risi, S., Lambrecht, L., Cook Jr, E. H., Leventhal, B. L., DiLavore, P. C., … Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. Journal of Autism and Developmental Disorders, 30(3), 205–223.

Lord, C., Rutter, M., DiLavore, P., & Risi, S. (2002). Autism diagnostic observation schedule: ADOS: Manual. Western Psychological Services Los Angeles, CA.

McCormack, J., McLeod, S., McAllister, L., & Harrison, L. J. (2009). A systematic review of the association between childhood speech impairment and participation across the lifespan. *International Journal of Speech-Language Pathology*, *11*(2), 155–170.

McGrath, L. M., Hutaff-Lee, C., Scott, A., Boada, R., Shriberg, L. D., & Pennington, B. F. (2008). Children with comorbid speech sound disorder and specific language impairment are at increased risk for attention-deficit/hyperactivity disorder. *Journal of Abnormal Child Psychology*, *36*(2), 151–163.

Miller, J., & Chapman, R. (2008). *Systematic Analysis of Language Transcripts*. Baltimore: University Park Press.

Muthén, L. K., & Muthén, B. O. (1998-2011). Mplus User's Guide. Sixth Edition. Los Angeles, CA: Muthén & Muthén.

Nahum-Shani, I., Qian, M., Almirall, D., Pelham, W. E., Gnagy, B., Fabiano, G., … Murphy, S. (2012). Experimental Design and Primary Data Analysis Methods for Comparing Adaptive Interventions. *Psychological Methods*, *17*(4). http://doi.org/10.1037/a0029372

Nelson, H. D., Nygren, P., Walker, M., & Panoscha, R. (2006). *Screening for Speech and Language Delay in Preschool Children*. Rockville (MD): Agency for Healthcare Research and Quality (US). Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK33412/

Paul, R. (1993). Patterns of development in late talkers: Preschool years. *Communication Disorders Quarterly*, *15*(1), 7–14.

Paul, R. (2000). Predicting outcomes of early expressive language delay: Ethical implications. *Speech and Language Impairments in Children: Causes, Characteristics, Intervention and Outcome*, 195–209.

Paul, R., & Jennings, P. (1992). Phonological behavior in toddlers with slow expressive language development. *Journal of Speech, Language, and Hearing Research*, *35*(1), 99–107.

Paul, R., & Shiffer, M. E. (1991). Communicative initiations in normal and late-talking toddlers. *Applied Psycholinguistics*, *12*(04), 419–431. http://doi.org/10.1017/S0142716400005853

R Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

Roberts, M. Y., & Kaiser, A. P. (2015). Early Intervention for Toddlers With Language Delays: A Randomized Controlled Trial. *Pediatrics*, peds.2014–2134. http://doi.org/10.1542/peds.2014-2134

Rowe, M. L., & Goldin-Meadow, S. (2009). Early gesture selectively predicts later language learning. *Developmental Science*, *12*(1), 182–187. http://doi.org/10.1111/j.1467-7687.2008.00764.x

Smith, V., Mirenda, P., & Zaidman-Zait, A. (2007). Predictors of expressive vocabulary growth in children with autism. *Journal of Speech, Language, and Hearing Research*, *50*(1), 149–160.

Stanton-Chapman, T. L., Justice, L. M., Skibbe, L. E., & Grant, S. L. (2007). Social and Behavioral Characteristics of Preschoolers With Specific Language Impairment. *Topics* *in Early Childhood Special Education*, *27*(2), 98–109. http://doi.org/10.1177/02711214070270020501

Steiger, J. H., & Lind, J. C. (1980, May). Statistically based tests for the number of common factors. In *annual meeting of the Psychometric Society, Iowa City, IA* (Vol. 758).

Stone, W. L., & Ousley, O. Y. (1997). STAT Manual: Screening tool for autism in two-year-olds. *Unpublished manuscript, Vanderbilt University*.

Thall, P. F., Logothetis, C., Pagliaro, L. C., Wen, S., Brown, M. A., Williams, D., & Millikan, R. E. (2007). Adaptive therapy for androgen-independent prostate cancer: a randomized selection trial of four regimens. *Journal of the National Cancer Institute*, *99*(21), 1613–1622. http://doi.org/10.1093/jnci/djm189

Thurm, A., Lord, C., Lee, L.-C., & Newschaffer, C. (2006). Predictors of Language Acquisition in Preschool Children with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *37*(9), 1721–1734. http://doi.org/10.1007/s10803-006-0300-1

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, *40*(6), 1245–1260.

Toth, K., Munson, J., Meltzoff, A. N., & Dawson, G. (2006). Early predictors of communication development in young children with autism spectrum disorder: Joint attention, imitation, and toy play. *Journal of Autism and Developmental Disorders*, *36*(8), 993–1005. http://doi.org/http://dx.doi.org.proxy.library.vanderbilt.edu/10.1007/s10803-006-0137-7

Ungerer, J. A., & Sigman, M. (1981). Symbolic play and language comprehension in autistic children. *American Academy of Child Psychiatry*, *20*, 318–337.

Wake, M., Tobin, S., Girolametto, L., Ukoumunne, O. C., Gold, L., Levickis, P., … others. (2011). Outcomes of population based language promotion for slow to talk toddlers at ages 2 and 3 years: Let's Learn Language cluster randomised controlled trial. *Bmj*, *343*.

Warren, S. F., Fey, M. E., & Yoder, P. J. (2007). Differential treatment intensity research: A missing link to creating optimally effective communication interventions. *Mental Retardation and Developmental Disabilities Research Reviews*, *13*(1), 70–77.

Watt, N., Wetherby, A., & Shumway, S. (2006). Prelinguistic Predictors of Language Outcome at 3 Years of Age. *Journal of Speech, Language, and Hearing Research*, *49*(6), 1224–1237. http://doi.org/10.1044/1092-4388(2006/088)

Yoder, P., & Stone, W. L. (2006). A randomized comparison of the effect of two prelinguistic communication interventions on the acquisition of spoken communication in preschoolers with ASD. *Journal of Speech, Language, and Hearing Research*, *49*(4), 698-711.

Yoder, P., & Compton, D. (2004). Identifying predictors of treatment response. *Mental Retardation and Developmental Disabilities Research Reviews*, *10*(3), 162–168. http://doi.org/10.1002/mrdd.20013

Yoder, P., & Symons, F. (2010). *Observational Measurement of Behavior* (1 edition). New York: Springer Publishing Company.

Yoder, P., Watson, L. R., & Lambert, W. (2015). Value-Added Predictors of Expressive and Receptive Language Growth in Initially Nonverbal Preschoolers with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 1–17. http://doi.org/10.1007/s10803-014-2286-4

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). *Preschool Language Scales, Fifth Edition (PLS^{TM}-5)*. San Antonio, TX: Pearson Education, Inc.

Zimmerman, Steiner, & Pond, R. E. (2009). *Technical Report: Preschool Language Scale-Fourth edition*. Pearson Education.