2016

Receptive verb knowledge in the second year of life: an eye-tracking study

BOSTON UNIVERSITY

SARGENT COLLEGE OF HEALTH AND REHABILITATION SCIENCES

Thesis

RECEPTIVE VERB KNOWLEDGE IN THE SECOND YEAR OF LIFE:

AN EYE-TRACKING STUDY

by

MATTHEW JAMES VALLEAU

B.A., Northeastern University, 2011
M.A., Boston University, 2014

Submitted in partial fulfillment of the

requirements for the degree of

Master of Science

2016

Approved by

First Reader ______________________________
Sudha Arunachalam, Ph.D.
Assistant Professor of Speech, Language & Hearing Sciences

Second Reader ______________________________
Kristine Strand, Ed.D., CCC-SLP
Clinical Associate Professor of Speech, Language & Hearing Sciences

Third Reader ______________________________
Karole Howland, Ph.D., CCC-SLP
Clinical Assistant Professor of Speech, Language & Hearing Sciences

DEDICATION

I would like to dedicate this work to my parents, Anne and Clint Valleau, for the unconditional love and support that have shown me all my life.

ACKNOWLEDGMENTS

I would like to thank the members of my thesis committee, Sudha Arunachalam, Kristine Strand and Karole Howland for their support over the past two years.

I would also like to thank the families and children that donated their time and effort to participate in this study, and all of the members of the Boston University Child Language Lab who tirelessly recruited, scheduled, organized and ran the present study, especially Leah Sheline, Shawn Dennis and Jessica Brough, as well as Ju Hong, for graciously processing the data several times.

Thank you to the National Institute of Health, for supporting this study (NIH K01DC013306).

Finally, I'd like to thank my family, friends and classmates, as well as the faculty and staff of the Boston University Department of Speech and Hearing Sciences for their continued support through my academic career and the future.

RECEPTIVE VERB KNOWLEDGE IN THE SECOND YEAR OF LIFE:

AN EYE-TRACKING STUDY

MATTHEW JAMES VALLEAU

ABSTRACT

The growth of a child's early vocabulary is one of the most salient indicators of progress in language development, but measuring a young child's comprehension of words is non-trivial. Parental checklists are prone to underestimation of a child's vocabulary (Houston-Price et al., 2007; Brady et al. 2014), so it may be that more direct measures, such as measuring a child's eye movements during comprehension, may provide a better assessment of children's vocabulary. Prior research has found relationships between gaze patterns and vocabulary development (Fernald et al. 2006), and the present exploratory study investigates these relationships with verbs, along with a number of methodological considerations. In addition, recent research supports the idea that verbs may differ in difficulty of acquisition based on word class, with manner verbs being easier to learn than result verbs (Horvath et al. 2015). The present study has two aims: 1) investigate the effect of dynamic stimuli on correlations with vocabulary scores and 2) experimentally investigate the notion that manner verbs are easier to learn than result verbs.

Forty children (Mean age = 22.97 months) were recruited for participation and shown a vocabulary test. While no significant correlations were found between vocabulary measures and accuracy and latency, several experimental measures proved to be related to vocabulary development, including fixation density and length of first fixation to the non-target. Additionally, results indicate that children knew the same number of manner and result verbs. Finally, these results could inform vocabulary tests using eye-tracking measures that specifically target verb knowledge.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

C .... Target (Correct)

I .... Non-target (Incorrect)

IPL .... Intermodal Preferential Looking Paradigm

INTRODUCTION

The growth of a child's early vocabulary is one of the most salient indicators of progress in language development. In order to learn a new word, a child must process the incoming linguistic input and identify the word's phonological form. Taking for granted the first two steps, she must then identify the referent of the new word, and store this pairing in memory in order to be able to recognize or produce the word later on (Gupta & Tisdale, 2009). This fast-mapping, as this process has come to be known, can be incredibly rapid (Carey, 1978), but has also been shown to create fragile links, with poor retention reported (Munro, Baker, McGregor, Docking and Arculi, 2012; Horst and Samuelson 2009).

The process for learning new words may seem simple, but is actually quite complex: the infant must notice the coincidence of the acoustic pattern in the linguistic stream, and map that onto a co-occurring referent in the world. Quine (1960) famously articulated the problem: imagine a speaker of a foreign language pointing to a rabbit and saying *gavagai*. What possible meanings could be proffered as the meaning of this unknown word? Perhaps rabbit, furry, or future-good-luck-charm. Since then, a number of different mechanisms have been posited as a solution to this seemingly impossible problem such as a variety of constraints or assumptions that an infant makes when encountering new nouns (Markman 1991). Although the process for nouns is quite complex, for verbs the undertaking is even more daunting. Whereas the referents of many nouns are concrete and stable in the world, verbs represent relations, and are inherently abstract (Gentner, 1978). One can point to a ball and identify it with the word;

identifying an action like kicking becomes more difficult as one must ascertain the correct relation between the kicker and the kickee. This conundrum is more clearly illustrated by pairs of verbs that could describe the same action such as pouring and filling. Furthermore, actions in the world, in contrast with objects, are transient, and may be labeled before, during, or after by their co-occurring linguistic construct, verbs, complicating identification by a young language learner (Gillette et al. 1999). Verbs are also more difficult than nouns to learn simply by observing the visual context in which they are discussed. Gillette et al. (1999) used silent videos and asked adults to name the verbs and nouns that a mother was saying at a particular instant in conversation with their child. Adults named far fewer verbs than nouns in this experiment, underscoring the fact that for verbs, the extralinguistic context is not enough to figure out what is being said, and that more information must be taken into consideration in order to learn verbs.

The above differences between verbs and nouns have repercussions for infants that are reflected in lexical development, specifically with regards to the composition of early lexicons and the relative ease with which nouns and verbs can be learned as shown in both corpus and experimental data. First, corpus data suggests that early lexicons are made up of a disproportionate number of nouns relative to verbs, compared to an adult's lexicon. This difference holds up cross-linguistically (Bornstein et al., 2004; Imai et al., 2008; Tardif, Gelman & Xu, 1999), suggesting that something about verbs is harder, even despite linguistic differences that may put verbs in different languages at different levels of salience. For example, languages like Mandarin Chinese allow speakers to omit nouns if the information is recoverable from the context; the effect of this is a higher percentage of verbs in the output, relative to nouns, when compared with English-language discourse, placing verbs at a higher level of prominence in Mandarin speech (Tardif, 1996). Further research on this noun/verb discrepancy has shown that the percentage of verbs in the total vocabulary takes some time to grow, with adult-like proportions of nouns to verbs achieved only by 30 months, a further indication that verbs are acquired more slowly than nouns (Fenson et al., 1994). Finally, in an experimental study, Childers & Tomasello (2002) taught children six novel nouns or six novel verbs, varying the presentation style: numerous repetitions over one day, or numerous repetitions over two weeks. They found that two year olds learned and "produced three times as many nouns as verbs," lending experimental support to the idea that verbs are more difficult to learn than nouns (p. 976).

While it is clear that verbs are more difficult to learn than nouns, it may also be the case that certain classes of verbs represent differing challenges to children, for example, manner and result verbs. Manner verbs are verbs that encode the way in which something is done (e.g., *run*), while result verbs encode the final outcome of the event (e.g., *break*) and research shows that the 'result' distinction (i.e., knowing that a verb requires a change in state) may not be a salient aspect of verb knowledge for children below the age of seven (Gentner, 1978). Recent research shows that typically developing children know more manner verbs than result verbs suggesting that some semantic classes of verbs may be more difficult for young learners (Horvath, Rescorla & Arunachalam, 2015; Gentner, 1978; Gropen et al., 1991; but see Behrend, 1990).

The ease with which a child learns new words has broader implications, as early vocabulary development, and verb vocabulary in particular, has been linked to a variety of language outcomes. In a large-scale, longitudinal study of 1,071 children from the ages 2 to 11, Lee (2011) found that vocabulary size and the number of verbs in a child's early vocabulary were strong predictors of a variety of language outcomes, among them phonological awareness, reading comprehension, and grammatical development. Marchman and Bates (1994) highlight that growth in the number of verbs in the early lexicon may drive grammatical development, supporting the critical mass hypothesis which states that the young learner must "must acquire a certain number of words, especially verbs, before progressing to learn the grammar of the language" (as cited in Lee, 2011). This further emphasizes the role that verb acquisition plays in early and later language development, and assessment of that lexical knowledge is an important aspect of understanding a child's overall language development.

Parental checklists are a common way to estimate a child's vocabulary, but studies have documented problems with using a checklist instead of a direct measure. For example, Houston-Price, Mather and Sakkalou (2007) used an Intermodal Preferential Looking paradigm (IPL), in which the child was shown two images and asked to look to one of them. If the child looks to the correct target, it can be inferred that the child knows what that word is; this is thus a more direct measure of the child's comprehension, as it measures the child's behaviors, and doesn't rely on report from a third party. Houston-Price et al. (2007) found that parents routinely underestimated their child's lexical development and it has been found that verbs are particularly susceptible to such underreporting, potentially due to parents attending more to nouns during language-centered activities and thus not noticing verbs their child had just produced (Tardif, Gelman and Xu, 1999).

A number of other studies have utilized this paradigm in order to assess concurrent validity for checklists and other measures of vocabulary. Styles and Plunkett (2009) tested infants' word knowledge using IPL and found that parents' estimations of their children's vocabulary were accurate on the item level, demonstrating that IPL can be used to ascertain lexical comprehension in children as young as eighteen months. Brady, Anderson, Hahn, Obermeier, and Kapa (2014) recently investigated the performance of children with autism on the Peabody Picture Vocabulary Test (PPVT, Dunn & Dunn 2007) in comparison with an IPL paradigm using an eye tracker. They found that children with autism and typical development both looked to the target longer when they knew the word already (as assessed by the PPVT). Although the children with autism and neurotypical children performed similarly for trials where they knew the word, there were significant differences between looking patterns in unknown trials: typically developing children seemed to understand words that the PPVT had indicated were unknown. The authors attribute this to a possible underestimation of several of the neurotypical subjects' vocabularies, as well as a failure to capture emergent vocabulary knowledge, which further supports the problem raised in Houston-Price et al. (2007). Furthermore, in a recent study, Venker et al. (2016) investigated fast mapping skills in children with autism (age 3.5 years), in relation to both current and later language development. They found that the ability to fast-map new vocabulary was related to receptive language development, with poor learners scoring significantly lower on the

Preschool Language Scale – 4th Edition both at age 3.5 years and two years later. This further highlights the importance of assessment of comprehension and vocabulary growth as a window into concurrent and later language development.

These papers show that IPL can be used to assess children's word knowledge in a variety of populations. Despite the importance of verb knowledge for language outcomes, none of these studies have investigated verb knowledge using IPL or eye-tracking. We will thus employ a similar paradigm to focus on verbs using dynamic scenes of actions employing an eyetracker. Using a video with a dynamic scene allows us to test children using much more realistic stimuli than are used in other vocabulary measures, which can be problematic. The PPVT, for example, uses static images to display actions, which some research suggests may be beyond the interpretive capabilities of young children (Cocking & McHale, 1981). Additionally, this paradigm does not require a motor response such as pointing or head turning, and thus can test children with motoric difficulties that extend beyond speech impairments. Using this paradigm thus avoids the problems associated with indirect measures of vocabulary such as parent report, and can be extended to a group of children that are harder to test, namely those who are younger, or impaired in some way that prevents them from pointing or speaking responses. Finally, utilizing an eyetracker provides us with data with a very high temporal resolution, allowing not only evaluation of comprehension by looking at the average amount of time spent looking at target, but tracking incremental processing of speech via eye movements (Fernald, Zangl, Portillo and Marchman, 2008). This is the main methodological difference between IPL and the current study: in IPL the data are averaged in a time window in order to state the overall preference of looking to one target, but we are tracking the child's gaze frame-by-frame over time, and thus can use a finer analysis than aggregating eye gaze patterns allows for.

The aims of the current study are twofold. The first aim is to investigate both established and exploratory predictors of verb knowledge related to eye-movement behaviors specifically with regards to static versus dynamic stimuli. We have chosen four variables related to participants' eye movements that may relate to their underlying processing or knowledge. The first variable is overall accuracy, which is defined as the proportion of time spent looking to the target to time spent looking elsewhere. The second variable is latency, which will be measured as the duration between the appearance of the two images and the child's first look to the target picture. Together, accuracy and latency have been shown to correlate with lexical and grammatical development in the second year of life, and thus have already been shown to be reliable indicators of vocabulary development (Fernald et al., 2006; Fernald et al., 2008). While this relationship has been upheld for nouns represented with static images, this has not been investigated for verbs using dynamic stimuli. We hypothesize that these relationships will be altered for dynamic stimuli, as dynamic scenes require a certain amount of time in order to perceive the event that is taking place and they also constantly draw attention due to movement. This hypothesis leads to a number of predictions. First, we expect to confirm prior research with nouns and static images, and predict that accuracy and latency for noun trials will correlate with lexical knowledge. Because dynamic scenes may require different amounts of time to fully appreciate the depicted event, we predict that overall accuracy, but not latency will correlate with lexical knowledge on verb trials. As prior literature has not looked at the distinction between dynamic and static stimuli, a number of other exploratory measures have been devised for the present study, primarily influenced by the literature surrounding how adults view complex scenes. The third variable we will measure is fixation density, or the ratio of the number of fixations to the non-target to the number of fixations to the target, and the fourth is the length of the first fixation to the target. These latter two measures have been shown to relate to semantic consistency in complex scene viewing, and may give insight into online processing during speech comprehension in the course of our task (Henderson, Weeks, & Hollingworth, 1999). Specifically, Henderson et al. (1999) showed adults scenes with semantically consistent and semantically inconsistent elements (e.g., consistent: a bar with a martini glass; inconsistent: a bar with a microscope). They found that participants fixated on semantically consistent targets fewer times than they did on semantically inconsistent targets. If we take the target scene to be "semantically consistent" with the auditory prompts, we predict that infants will have more fixations to the non-target than to the target. Thus, the proportion of fixations to the non-target to fixations to the target will correlate with vocabulary for both noun and verb trials. Henderson et al. (1999) further found that initial saccades to semantically consistent target objects were shorter than those to semantically inconsistent targets. Using the same definition of semantic consistency above, we would expect to see shorter initial fixations to the target and longer initial fixations to the non-target, and predict that the length of the first fixation to the non-target will correlate with lexical knowledge.

The second aim of the present study is to experimentally investigate the recent finding by Horvath et al. (2015) that typically developing children know more manner verbs than result verbs. We hypothesize that result verbs are more difficult than manner verbs and therefore predict that children will know more manner verbs than result verbs, as ascertained by accuracy during those trials. Accuracy was chosen, as opposed to both accuracy and latency, because accuracy measures overall looking preference which can be used to infer a child's vocabulary knowledge, whereas latency only shows how quickly they looked to the target, and after the target is established, children have a fifty-fifty chance of looking to the target immediately given the current study's stimulus design (see below for full descriptions).

In summary, we have two hypotheses: (1) that dynamic stimuli encourage different eye gaze patterns than static stimuli, thus altering previously established relationships between eye gaze measures of lexical knowledge; and (2) that result verbs are more difficult for children to learn than manner verbs and that this will be reflected in the number of each type of verb in present in early lexicons. With these in mind, we make the following predictions:

1. Latency and accuracy in noun trials (static images) will correlate with lexical knowledge
2. Accuracy, but not latency, in verb trials will correlate with vocabulary development
3. Proportion of the number fixations to the non-target to the number of fixations to the target in both noun and verb trials will correlate with lexical knowledge

4. Length of the first fixation to the non-target will correlate with vocabulary development in both noun and verb trials
5. Children will know more manner verbs than result verbs, as measured with overall looking time to targets in trials with those types of verbs.

METHODS

Participants

In order to be included in the present study, potential participants needed to meet the following criteria:

1) be between the ages of 22.0 and 24.9 months, as this is right in the middle of a period of "smoothly accelerating exponential" vocabulary growth, and in particular growth in the number of verbs in the child's lexicon relative to the number of nouns (Fenson et al. 1994),

2) have had a full-term gestational period (37+ weeks),

3) have no more than 30% exposure to languages other than English, and

4) have no history of speech, language or hearing disorders, or developmental delays.

Participants who met the necessary inclusionary criteria for the study were randomly assigned to one of four possible groups, A, B, C, and D, specifying which version of the stimuli were used (see the Apparatus and Stimuli section for a full description of the groups). Groups A and B, and groups C, D, were matched with each other on age (within 1 month), gender, and vocabulary scores (within 8 points) on the

MacArthur-Bates Short Form Vocabulary Checklist: Level II (MCDI-2; Fenson, 2007). This matching was done to control for chronological age, vocabulary size, and gender to avoid potential confounds.

Sixty-six children were recruited for the experiment. Ten of those were excluded: five due to poor eye tracking; one due to parental interference during the test; one due to an uncooperative child; two who were in early intervention for language delays; and one with a suspected expressive language delay. From the remaining eligible children, a sample of forty was selected, ten from each group, based on the matching variables described above, a sample size which is consistent with prior research done using this paradigm (Houston-Price, Mather & Sakkalou, 2007; Brady et al., 2014). The sample was made up of twenty males and twenty females, with an average age in months of 22.97 (sd = .69), and an average MCDI score of 56.3 (sd = 22.5). Table 1 shows each participant's group, sex, age, and MCDI score. Groups did not differ significantly based in age or vocabulary development as measured by the MCDI, and were thus collapsed into one cohort for the analyses below.

Apparatus and Stimuli

The data were collected using Tobii T60 XL Eye Tracker, which has a 24" monitor and built in speakers with which the stimuli are presented. The eyetracker requires the participant to be between 50 and 80 cm from the screen in order to track eye movements accurately, and has a maximum latency of 33 milliseconds (ms), and tracking accuracy of .5 degrees.

The test stimuli consist of a video with twenty-five trials of nouns or verbs, chosen by Konishi et al. (2012) for their high imageability as well as their likelihood to be present in toddlers' vocabularies (see Appendix 1 for a diagram of the timing in one trial). The verbs include both manner verbs (e.g., dance) as well as result verbs (break). The visual stimuli used were first developed by Konishi et al. (2012) to assess children's vocabulary knowledge in a pointing task and were adapted for use in an eye-tracking task by adding the auditory prompts to the videos, as well as determining the precise timing of the presentation. For each trial, the child is presented with two images, one at a time, followed by a joint presentation of both. These images disappear, and the child is asked to find one of them (the target). Both of the images then reappear, and the child is again asked to find the target.

Noun trials present the child with two static photographs of objects, while verb trials present the child with two dynamic videos in which actions are taking place with live actors. The trials are counterbalanced for which side is presented first (right versus left), and which side is the target (right versus left), as well as for the order of the trials (forward or backward). Full descriptions of counterbalancing for trials can be found in Appendix 2.

In addition, there are four groups (A, B, C, and D, discussed above) using the same videos, but varying the order of the trials, as well as which picture is the target. In other words, the videos across all participants are identical, but the audio is varied in order to target the other picture, controlling for the visual input as a possible confound. For example, if one element of a scene is more salient for some reason (eg. color) and thus draws more attention because of this feature and not due to the linguistic input, this effect will be the same across groups, eliminating that feature as a source of possible group differences. In addition to controlling for the visual stimuli, there are two other reasons for this experimental design. The first reason is to avoid showing participants the same video twice in order to eliminate the possibility of a child choosing a target by excluding the target they had identified previously. Furthermore, this design allows us to reduce the amount of time for each participant, while maximizing the number and variety of trials we can use.

Procedure

Upon arrival to the laboratory, the participant's caregiver is given a consent form to sign, as well as the MCDI-2 and an additional vocabulary questionnaire to fill out. The child is allowed to play with various toys in the lab while the caregiver finishes the necessary paperwork. After this is complete, the participant and caregiver are brought in to the room with the Tobii Eye tracker. The child is either placed in a car seat in front of the eye tracker, or placed on the caregiver's lap. If the caregiver holds the child, the caregiver is asked to wear an eye mask in order for their eye movements to not be captured and so that the caregiver doesn't influence not only the child's gaze patterns, but also potentially her body movements or position. One experimenter runs the Tobii software from behind a curtain, and the other directs the child's attention to the screen.

Data Analysis

The raw data, consisting only of where a child was looking on the screen at a given time point in the showing of a video, must be processed before analysis. This preprocessing uses a script in R developed to identify all the trials and sub-phases within those trials based on their time stamp. Appendix 1 shows a schematic version of one trial, including the duration and name of each sub-phase. The target is established (e.g., auditory stimulus: “Can you show me the cookie?”), while the screen had a generic place holder in the center (“Star” sub-phase), meant to draw participant’s gaze away from either side, immediately followed by the “Response” sub-phase. As such, only the “Response” sub-phase was analyzed. After pre-processing of the data was complete, a variety of measures were then calculated, including measures established in the literature (accuracy and latency) as well as our exploratory measures (fixation density and length of first fixation), which required further processing in order to identify fixations in the raw data. The algorithm that was developed is detailed below.

Calculation of measures.

Two measures, accuracy and latency, have been shown to be robust predictors of a child’s vocabulary development, and as such, were selected both as an attempt to replicate prior results with the nouns, as well as seek a relationship with verbs using dynamic videos. Please note that these calculations were done on the raw data, that is, without identifying fixations, which is discussed below for the experimental measures.

Accuracy.

In prior literature (e.g., Fernald et al. 2006) accuracy is calculated as the ratio of time spent looking to the target to time spent looking to the non-target, excluding looks away from the screen. However, having a ratio of this nature proved difficult during analysis when the denominator (time spent looking to the non-target) was zero; to solve this, a percentage of looks to the target out of all looks to either the target or the non-target was calculated, avoiding the problem of dividing by zero. In addition, some studies have calculated accuracy as a percentage of looks to the target out of the entire time window, thus potentially including looks off-screen. This method of computing accuracy was also included, yielding three methods for calculating accuracy:

1. The ratio of time spent looking to the target to time spent looking to the non-target
2. The percentage of time spent looking to the target out of time spent looking to either the target or non-target
3. The percentage of time spent looking to target out of the entire window

Latency.

In addition to accuracy, latency is a measure that has shown a robust relationship with vocabulary development. Latency was measured as the difference between the time at which participants initially looked to the target and the time when the video was presented for the response phase of the trial.

I-VT Algorithm and calculation of associated measures.

In addition to the above measures, the following measures were designed as an attempt to capture the effects of dynamic stimuli on participant's gaze patterns, in an effort to find a relationship between these measures and overall vocabulary development. Calculation of these measures entailed processing the data further and identifying fixations and saccades reliably. The algorithm that was developed is described below, as well as the calculations for each of the experimental measures.

Algorithm Description.

Velocity threshold identification (I-VT) was selected for identification of fixations and saccades due to its robustness and ease of implementation (Salvucci & Goldberg, 2000). The overall algorithm functionality in pseudocode is as follows:

1. Calculate point-to-point velocities for each point in the protocol
2. Label each point below velocity threshold as a fixation point, otherwise as a saccade point
3. Collapse consecutive fixation points into fixation groups, removing saccade points
4. Map each fixation group to a fixation at the centroid of its points
5. Return fixations (Salvucci & Goldberg, 2000)

The first step in this algorithm is non-trivial, because the velocity of an eye movement is angular in nature, due to the rotational movement of the eye. This entailed mapping the location of each eye in relation to the screen (width, height, and distance) and calculating a distance to the screen coordinates at that time point, and comparing this to the next point. The Tobii system described above automatically generates the eye positions and gaze positions in a suitable coordinate system, but the algorithm to calculate distance and angular velocity was written in the Python programming language for the current study, with the threshold velocity 100 degrees/second, a standard threshold for fixation speed (Salvucci & Goldberg, 2000). The output of this algorithm was a list of fixations for each participant, coded for the order in which the fixation occurred, the duration of the fixation, and the centroid of the fixation, which was used to calculated whether the fixation occurred on the target, the non-target, or neither. This list was then used to compute the experimental measures below.

- Fixation density: the proportion of the number of fixations to the non-target to the number of fixations to the target
- Length of First Fixation: the length of the first fixations to the target, as well as the length of the first fixation to the non-target using the duration computed during fixation identification.

Time Windows.

Variability may be seen in gaze patterns when first presented with a question, versus a few seconds after, when attention may have begun to wane; in order to capture an accurate picture, it is necessary to bound analyses to a particular time window within a trial. A number of time windows were investigated in the present study. The window that has been used in previous studies showing relationship between accuracy and latency and vocabulary development was between 300ms and 1800ms during the response phase, as 300ms are needed to program and launch an eye movement (Fernald et al, 1998, Fernald et al. 2006). However, a number of methodological differences exist between those studies and the present study. In those studies, the participant hears the target during the response phase, unlike the present study, where participants are presented with the question (e.g., “Where is the cookie?”) before the visuals are onscreen. Additionally, the present study had a central fixation point, whereas prior studies did not. These methodological differences could lead to differences in eye-movements, if for instance children remember where the target is, or they happen to be faster processors. Taking this into account, a widened window of 0ms-1800ms was also investigated. Finally, because the present study's stimuli have dynamic videos, children might need more time to fully process a given event that is happening in the scene. A preliminary analysis showed that for a given video, children who were asked to find different verb targets (e.g., either "feeding" or "hugging") showed divergence in their looks the their respective targets around 500ms, and maximal differences around 2000ms after the sub-phase onset, whereas a similar pattern over a much shortened time frame was observed for noun targets, supporting a wider window of analysis for dynamic stimuli. A third window of 0ms-2000ms was thus selected in order to capture gaze patterns both at the very initial stages of viewing as well as those up until maximal divergence for the dynamic stimuli. Finally, for latency measures, two windows were used, one with the entire sub-phase included for analysis, and one from 0-3000ms. This second window served to exclude outliers in the data, namely participants who for took an excessive amount of time to look to the target (for example, due to distraction in the room, or not knowing the correct target, among other potential reasons).

Area of Interest Coding.

Finally, the raw gaze data (or fixation data) represent a point on the screen which must be mapped onto the content of the screen. Two ways to do this were investigated. Using an eye-tracker allows for a very fine spatial resolution for the localization of an infant's gaze on the screen, and it is with this in mind that narrow areas of interest were defined, localized only to the video itself, not the whitespace around the video (see Appendix 3 for a schematized version). However, other factors such as fixation drift could cause the coordinates to cross the narrow delineation between the video and the whitespace around it, and this could lead to a classification outside of the area of interest given the fixation identification algorithm. Additionally, many previous studies did not use eye-tracking and instead used human coders to categorize infants' gaze as either to the left or the right, representing a broad area of interest. Since the current study aimed in part to replicate findings from these studies, this second method to calculate areas of interest was also used in the analyses below.

RESULTS AND DISCUSSION

A series of analyses were applied to the data, with the specifications below, in order to answer the present study's research questions. Correlations were used in order to ascertain the effects of dynamic stimuli on gaze patterns and establish relationships between potential eye gaze measures and vocabulary knowledge. In addition, recent findings regarding the relative difficulty of manner and results verbs were investigated.

Correlational Analyses

Accuracy and Latency

Accuracy and latency have been shown to be related to overall vocabulary development in previous studies and as such, these two measures were investigated first. Correlation was used to examine the relationship, if any, of these measures, as calculated above, with both the total MCDI score, as well as with the noun and verb scores of the MCDI. Recall the predictions outlined above, namely that a) accuracy and latency would correlate with MCDI for noun trials and that b) accuracy, but not latency would not correlate with MCDI for verbs, given the dynamic nature of the stimuli. These will now be considered in turn.

Accuracy

Table 2 shows the Pearson correlation coefficients for accuracy and vocabulary scores, broken up by AOI coding method, window, trial word class, and MCDI versus the scores for nouns or verbs on the MCDI, similarly to the breakdown for latency above. Narrow AOI coding was used initially, counting only a coordinate within the small stimulus video (see Appendix 3 for examples of the AOI coding styles). This coding revealed no significant relationship between accuracy and vocabulary development. Again, methodological differences in analysis could explain this failure to replicate prior findings, this time with AOI coding, as prior studies used human coders to identify looks to the target and non-target. With this course-grained approach, it is impossible to tell the difference between a look to the smaller video, or a look to the whitespace outside the video. Broad AOI coding was then used, splitting the screen down the middle in order to simulate this coding process. This method of AOI coding revealed subtle trends towards significant, with weak positive correlations for accuracy on noun trials (percentage of looks to the target out of looks to the target and non-target) with both MCDI and the noun score on the MCDI, evident across all three windows (e.g., window 0-1800ms, $r(38) = .27$, $p = 0.063$). This trend towards significance may have been hampered by the relatively small number of noun trials each child was exposed to, and it is possible that with an increase in the number of trials, a statistically significant relationship would be found. This finding supports a broad AOI coding scheme, and partially supports findings from past research on the relationship between accuracy and vocabulary development with regards to nouns. Furthermore, it highlights one method of accuracy calculation (percentage of looks to the target out of looks to the target and non-target) as being potentially relevant for future studies. However, contrary to the current study's predictions, no significant relationship was found between accuracy and vocabulary development for verb trials. This may be due to the difference between static versus dynamic stimuli drawing the gaze in a particular way, or to the nature of nouns and verbs; the present study cannot tease these two alternatives apart. However, it may be the case that accuracy is not a good measure for either dynamic stimuli or verbs, which may require a more complex measure, which will follow.

Correlations between accuracy on noun trials and accuracy on verb trials were computed in order draw a bridge between the noun and verb trials. Only one method of accuracy calculation showed significant weak-moderate, positive correlations between noun accuracy and verb accuracy: percentage of the total time spent looking to the target out of the time of the entire window. This relationship was stable across time windows and AOI coding methods (e.g., window 0-1800ms, $r(38) = 0.51$, $p < .001$). This relationship showed that children who tended to be accurate on noun trials, also tended to be accurate on verb trials, or perhaps an indicator that overall vocabulary development increases across word classes at this age.

Latency

Table 3 shows the Pearson correlation coefficients for these analyses, broken out by AOI coding type, window, word class and which vocabulary measure was used (either the full MCDI, or MCDI for nouns and verbs separately). No significant correlation was found between latency and vocabulary for nouns or verbs for any of the windows or AOI coding schemes. While lacking a correlation for nouns is contrary to the present study's predictions, given prior literature, the lack of correlation for verbs was predicted. However, given the inability to replicate prior findings, the ability to interpret this finding as supporting the hypothesis is limited. A number of methodological differences could cause the difference between the present study's findings and those of prior literature, with the primary factor being choice of targets, but also the number of trials per target. Fernald et al. (2006), which showed the correlations that the present study attempted to replicated, drew heavily from the methodology of an earlier study that looked comprehension at three different ages, and thus needed tokens to be easily recognizable to all ages (Fernald et al. 1998). This key difference, selecting targets which 15 month-old children are almost certainly likely to know (e.g., "baby", "doggy"), versus lower frequency targets that toddlers may or may not know (e.g., "crab" or verbs like "march"). Furthermore, each target in Fernald et al. (2006) was presented six times over the course of the videos, compared with the present study's one presentation per target. These two differences represent large methodological differences between past studies and the present study, and could very well have contributed to the failure to support past findings relating latency and vocabulary development.

In addition to correlations between latency and MCDI, correlations were run relating average latency in noun trials to average latency in verb trials, which yielded significant results: when taking into account the full response sub-phase, latency for nouns and latency for verbs showed a moderate positive correlation (r(38) = .46, p = .002), showing that children who were fast on noun trials were likewise fast on verb trials, and vice versa.

Eye gaze measures.

In addition to investigating previously established measures of processing efficiency, the present study used experimental measures to research similar relationships between viewing patterns and vocabulary development. These measures, instead of using the raw data as has been done for latency and accuracy, used fixations identified in the raw data as the basis for measurement of the two main variables below, fixation density and length of first fixation to the target and non-target.

Table 3 shows the Pearson correlation coefficients for all of the experimental measures and vocabulary measures, broken up by time window used and word class; narrow AOI coding was abandoned due to the findings above which showed prior relationships for accuracy only with broad AOI coding. Recall that the current study predicted that fixation density (number of fixations to non-target:number of fixations to target) would correlate with vocabulary for both nouns and verbs. Indeed, there is a significant relationship for both word classes, but with two important distinctions between them: for nouns, there is a weak negative relationship with vocabulary scores (e.g., window 300-1800ms, r(38) = -0.37, p=0.018), and with verbs there is a moderate positive correlation with vocabulary score (e.g., window 300-1800, r(38) = 0.44, p=0.004). There is also a difference in which time windows show this relationship: for nouns, the relationship appears in only the shorter windows, 0-1800ms, and 300-1800ms, while for verbs the relationship is stable across all three windows, including the longer window, 0-2000ms. This highlights the importance of choosing an appropriate time window depending on the stimulus type (static/dynamic) or word class (noun/verb), with a shorter window for static noun trials necessary in order to tease out the relationship between fixation density and vocabulary scores. In order to tease apart the reason these relationships are in opposite directions, one must consider the calculation method used: number of fixations to the target divided by the number of fixations to the non-target. These two measures were both calculated and correlated with vocabulary scores, and showed similar trends. Number of fixations to the target seems to be more relevant in noun trials, with marginal significance for the window 0-1800ms, when looking at fixation density to total MCDI score ($r(38) = 0.28$, $p=0.079$). Conversely, the number of fixations to the non-target in verb trials seems to be a more relevant measure, with weak positive correlations noted again across windows (e.g., window 300-1800ms, $r(38) = 0.33$, $p=0.034$). This differential pattern of relations between the number of fixations to either the target or non-target, may relate to the nature of the stimuli: a non-target dynamic scene may require a greater assessment of all the moving parts before a determination of the event type can be made, while a static scene may be more quickly appreciated. Again, the present study is not able to tease apart the relative effects of word class and stimulus type.

Finally, the last experimental measure that was investigated is the length of the first fixation to either the target or non-target. It was predicted that the length of the first fixation to the non-target would correlate positively with MCDI, for both nouns and verbs. Table 4 shows the Pearson correlation coefficients for these analyses. No significant relationship was found for noun trials, for either the length of the first fixation to the target or non-target. For verbs, however, a significant relationship was found, but not quite as predicted. Similar to fixation density, the fixation to the non-target seems to be more relevant when relating these measures to vocabulary development. Specifically, a moderate negative correlation was found (e.g., window 0-2000ms, $r(38) = -0.42$, $p=0.006$). This relationship was stable across with the full score of the MCDI and the verb score of the MCDI, perhaps illustrating that this relationship is verb specific in nature. Crucially, the relationship was only significant for the wider time windows (0-1800ms and 0-2000ms), reinforcing the idea that dynamic stimuli take a longer time to process, and that crucial information can be found in the first instants of viewing a complex scene. For verbs in particular, these two measures, fixation density and the length of the first fixation, with moderate negative correlations to vocabulary knowledge, paint a picture of the processing of dynamic scenes with verbal auditory input: children with larger vocabularies tend to look around the non-target event, perhaps confirming that it is not the target, while children with smaller vocabularies (and a smaller number of verbs in their vocabulary) tend to fixate in one position of the non-target event longer. Of course, with correlational analyses, correlation does not indicate causation, so it is impossible to tell from these analyses which came first: the larger vocabularies or the ability to quickly scan a dynamic event and decide on its relevance.

Manner and Result Verbs

In addition to the methodological investigation detailed above, the present study had a theoretical inquiry into differences in the acquisition of manner and result verbs. Specifically, given other research suggesting that result verbs may be more difficult than manner verbs for children to acquire (Gentner, 1978; Gropen et al., 1991; Horvath et al., 2015), it was hypothesized that this difference would show itself in an experimental context by children knowing a smaller percentage of the result verbs they were asked to identify in the present study. Knowledge of a particular verb was measured in terms of the accuracy computations detailed above, where an accuracy above 50%, or a proportion of looks to the target above 1:1 was regarded as correct. This was then averaged for each participant, and a paired t-test was used to assess the difference between the two means, which failed to reject the null hypothesis for any method of calculating accuracy (see Table 5 for t-test results). Figure 2 shows the average score for each group of verbs, along with standard deviations, showing a large overlap in the ranges of percentage correct for each verb type. This result is in contradiction with the current study's hypothesis that there would be a difference, with more manner verbs known than result verbs; however there are two methodological concerns that may influence interpretation of these data: sample size and potential order effects. Because the present study was designed to investigate both verbs and nouns, there are thus a limited number of trials for each word class. Additionally, the manner verbs represent a higher proportion of the verbs, with double the number of manner verbs represented (12 manner verbs versus 6 result verbs). Furthermore, Horvath et al. (2015) used parental checklists, which have been shown to be problematic when assessing comprehension, and the measure they used accounted only for expressive vocabulary. Thus, this discrepancy may be the result of a difference between comprehension and production of the two classes of verbs.

There are also potential order effects present in the sample, which could not have been predicted beforehand. One way ANOVAs were used to investigate group differences in accuracy for manner and result verbs, showing a significant effect of group on accuracy for manner verbs [$F(36, 3)=2.58$, $p= .0366$]. Post hoc comparisons using the Tukey HSD test indicated that the mean score for the B group was significantly lower than the C group (difference of means = -16.9%, $p = .0204$), illustrated in Figure 3. This is particularly surprising, as groups B and C differ only in the order of the targets presented, but otherwise saw the same videos and were asked to locate same targets. Given this equivalence in methods, an order effect of the stimuli, or randomness in the sample could be the cause of this effect.

CONCLUSION

The present study addressed two main questions. First, what are the effects of dynamic stimuli on gaze patterns, and are there measurements that can relate those patterns to overall vocabulary development, as has been done with accuracy and latency? Second, is the potential difference in difficulty between manner and result verbs illustrated in an experimental setting?

With regards to the first question, a number of measures were developed and tested. Specifically, fixation density (the proportion of the number of fixations to the non-target to the number of fixations to the target) and length of first fixation to the target were predicted to correlate with both nouns and verbs. Fixation density was in fact related with overall vocabulary development for both nouns and verbs, but for nouns the relationship was a negative correlation, while with verbs the relationship was positive. The underlying calculation was thus investigated, and a differential importance of fixations to the target and fixations to the non-target was found for nouns and verbs, respectively. Further, this differential relationship was supported in the length of first fixation, which showed positive correlations (trending towards significance) for nouns for the length of the first fixation to the target, but a positive correlation for verbs and vocabulary development when looking at the length of the first fixation to the non-target. These two findings together illustrate a big picture of children's processing: with increasing vocabulary there is more scanning of the non-target dynamic stimuli (leading to an increased fixation density, and a positive correlation with MCDI), while with decreasing vocabulary scores we see longer fixation on one point of the non-target, and less scanning of the non-target. Methodologically, the findings above support using different time windows for noun trials with static stimuli and verb trials with dynamic stimuli: smaller windows for nouns (0-1800ms and 300-1800ms) were more informative and longer windows for verbs (0-1800ms, 0-2000ms) were more informative. Additionally, broad AOI coding seems to be more informative when identifying fixations from the raw data and using the above measures of gaze patterns with dynamic stimuli.

With regards to the second question, the prediction was that children would know more manner verbs than result verbs. This was not supported by the current study, which found no difference in the percentage of known verbs of each class. While this result may be due to limitations in sample size, it may also show an interesting difference between expressive and receptive language, as Horvath et al. (2015) was taken from checklist data using what the child has produced, not what they understand.

There are a number of limitations of the current study, the first of which is the difference between stimulus type and word class. As all noun trials were static and all verb trials were dynamic, the ability to draw conclusions between is limited, and the relationship between the various measures that were devised may be difficult to interpret for static stimuli. Additionally, the types of verbs depicted favored manner verbs, lessening the strength of the conclusion that can be made with regards to the first research question. Finally, the failure to replicate previous findings is problematic, but given methodological differences between prior literature and the present study, it is not unreasonable that previous findings for comprehension were not shown in these analyses.

Future directions for this line of work include using more advanced statistical analyses such as regression to see if a combination of those measures above can give a greater picture of the relationship between gaze patterns and vocabulary development; such a tool could have more predictive power and may allow clinicians to appreciate a child’s vocabulary with a higher degree of certainty than correlational analyses allow. Additionally, exploring this approach with other populations should be considered a priority; we currently have data using these stimuli with children with autism, a group of children whose receptive language may be routinely underestimated (e.g., Swensen et al. 2007); the methods outlined above may provide a more accurate window into the language development of children with autism and enhance our understanding of autism both at a group and at an individual level.

Table 1: *Participants' ages and scores on MCDI by group*

	A (N=10)				B (N=10)			
	Age (months)	MCDI	MCDI-Nouns	MCDI-Verbs	Age (months)	MCDI	MCDI-Nouns	MCDI-Verbs
	M: 23.19	M: 54.7	M: 32.4	M: 4.4	M: 23.29	M: 53.6	M: 31.6	M: 4.2
	SD: 0.71	SD: 17.58	SD: 9.56	SD: 3.6	SD: 0.82	SD: 16.17	SD: 10.09	SD: 2.74
	24.47	73	44	7	23.95	70	41	7
	22.83	44	31	0	22.93	47	28	3
Female	23.88	64	42	6	24.47	57	37	4
	23.13	32	21	0	23.06	28	11	0
	23.29	65	35	9	22.5	66	37	6
	23.68	35	22	1	24	37	27	1
	23.36	80	42	9	24.37	74	43	8
Male	22.73	32	18	1	22.5	35	20	2
	22.24	62	39	6	22.34	55	34	4
	22.24	60	30	5	22.76	67	38	7

	C (N=10)				D (N=10)			
	Age (months)	MCDI	MCDI-Nouns	MCDI-Verbs	Age (months)	MCDI	MCDI-Nouns	MCDI-Verbs
	M: 22.78	M: 57.8	M: 31.6	M: 5.2	M: 22.66	M: 59.1	M: 30.7	M: 6
	SD: 0.66	SD: 29.35	SD: 14.84	SD: 5.29	SD: 0.37	SD: 27.25	SD: 13.12	SD: 6.15
	22.89	38	24	0	22.6	67	29	3
	22.11	71	37	7	22.57	75	43	15
Female	24.28	71	42	6	22.2	65	38	5
	23.06	99	49	14	23.55	100	50	15
	22.83	77	46	6	22.6	56	20	0
	22.76	22	11	1	22.57	21	18	3
	23.22	23	12	0	22.47	28	14	1
Male	22.17	99	49	14	22.86	98	48	14
	22.17	44	26	2	22.34	44	26	3
	22.34	34	20	2	22.83	37	21	1

Table 2: *Pearson correlation coefficients between average accuracy and vocabulary scores*

		AOI coding:	Narrow			Broad		
		Window:	0-1800ms	0-2000ms	300-1800ms	0-1800ms	0-2000ms	300-1800ms
Trial type	accuracy calculation method	measure						
Noun	ratio (C:I)	MCDI	-0.01	0.05	-0.00	0.08	0.08	-0.01
Noun	% (C/C+I)	MCDI	0.08	0.08	0.10	0.29†	0.29†	0.30†
Noun	% (C/total time)	MCDI	0.02	0.00	0.03	0.15	0.15	0.14
Verb	ratio (C:I)	MCDI	0.26	0.09	0.22	0.08	-0.03	0.02
Verb	% (C/C+I)	MCDI	0.15	0.15	0.15	0.01	0.01	0.05
Verb	% (C/total time)	MCDI	0.14	0.14	0.15	0.08	0.07	0.07
Noun	ratio (C:I)	MCDI-Nouns	0.02	0.11	0.05	0.09	0.09	0.01
Noun	% (C/C+I)	MCDI-Nouns	0.12	0.12	0.13	0.27†	0.27†	0.28†
Noun	% (C/total time)	MCDI-Nouns	0.04	0.02	0.05	0.14	0.14	0.14
Verb	ratio (C:I)	MCDI-Verbs	0.29†	0.10	0.25	0.16	0.08	0.10
Verb	% (C/C+I)	MCDI-Verbs	0.21	0.21	0.21	0.09	0.08	0.12
Verb	% (C/total time)	MCDI-Verbs	0.18	0.18	0.19	0.12	0.09	0.09
	ratio (C:I)-nouns	ratio (C:I)-verbs	0.16	0.00	0.06	0.21	0.16	-0.07
	% (C/C+I)-nouns	% (C/C+I)-verbs	0.12	0.02	0.12	0.02	0.01	-0.05
	% (C/total time)-nouns	% (C/total time)-verbs	0.51***	0.49**	0.50***	0.41**	0.41**	0.38*

note: ***: p <.001; **: p<.01; *: p<.05

†: .05 < p < .1

C refers to target (*correct*), I refers to non-target (*incorrect*)

Table 3: *Pearson correlation coefficients between average latency and vocabulary scores*

		AOI coding:	Narrow		Broad	
		Window:	full subphase	0-3000ms	full subphase	0-3000ms
Trial type	*x*	*y*				
Noun	Latency	MCDI	0.03	-0.06	0.03	-0.06
Verb	Latency	MCDI	-0.16	-0.10	-0.16	-0.10
Noun	Latency	MCDI-Nouns	0.01	-0.02	0.01	-0.02
Verb	Latency	MCDI-Nouns	-0.21	-0.17	-0.21	-0.17
	Latency-Nouns	Latency-Verbs	0.46**	0.26†	0.46**	0.26†

note: ***: $p < .001$; **: $p < .01$; *: $p < .05$

†: $.05 < p < .1$

C refers to target (*correct*), I refers to non-target (*incorrect*)

The measure latency is restricted to trials of the type labeled, except where noted otherwise

Table 4: *Pearson correlation coefficients between exploratory gaze pattern measures and vocabulary scores*

			AOI coding:	Broad		
			Window:	0-1800ms	0-2000ms	300-1800ms
Trial type	x	y				
Noun	Number fixations to C	MCDI		0.28†	0.25	0.21
Noun	Number fixations to I	MCDI		-0.12	-0.13	-0.11
Noun	Length first fixation to C	MCDI		-0.00	-0.00	-0.04
Noun	Length first fixation to I	MCDI		0.05	0.05	-0.00
Noun	Fixation density (I:C)	MCDI		-0.34*	-0.25	-0.37*
Verb	Number fixations to C	MCDI		0.24	0.24	0.21
Verb	Number fixations to I	MCDI		0.36*	0.34*	0.33*
Verb	Length first fixation to C	MCDI		-0.06	-0.07	-0.04
Verb	Length first fixation to I	MCDI		-0.39*	-0.42**	-0.26
Verb	Fixation density (I:C)	MCDI		0.42**	0.46**	0.44**
Noun	Number fixations to C	MCDI-Nouns		0.22	0.20	0.14
Noun	Number fixations to I	MCDI-Nouns		-0.08	-0.09	-0.11
Noun	Length first fixation to C	MCDI-Nouns		0.01	0.01	-0.01
Noun	Length first fixation to I	MCDI-Nouns		0.10	0.10	0.01
Noun	Fixation density (I:C)	MCDI-Nouns		-0.23	-0.14	-0.32*
Verb	Number fixations to C	MCDI-Verbs		0.28†	0.28†	0.25
Verb	Number fixations to I	MCDI-Verbs		0.37*	0.36*	0.36*
Verb	Length first fixation to C	MCDI-Verbs		-0.07	-0.08	-0.10
Verb	Length first fixation to I	MCDI-Verbs		-0.40*	-0.41**	-0.23
Verb	Fixation density (I:C)	MCDI-Verbs		0.45**	0.48**	0.47**

note: ***: $p < .001$; **: $p < .01$; *: $p < .05$

†: $.05 < p < .1$

C refers to target (*correct*), I refers to non-target (*incorrect*)

Table 5: *Paired t-tests between accuracy of manner and result verbs*

Window:	0-1800ms			0-2000ms			300-1800ms		
accuracy calculation method	mean of differences (manner-result)	t	p	mean of differences (manner-result)	t	p	mean of differences (manner-result)	t	p
ratio (C:I)	0.007225	0.2274	0.8213	0.0267625	0.8545	0.398	-0.0171825	-0.4914	0.6259
% (C/C+I)	0.0104475	0.3431	0.7334	0.0134275	0.4533	0.6528	0.004605	0.1530	0.8792
% (C/total time)	0.0003575	0.0101	0.9920	0.0292225	0.8522	0.3993	0.01792	0.5483	0.5866

note: C refers to target (*correct*), I refers to non-target (*incorrect*)

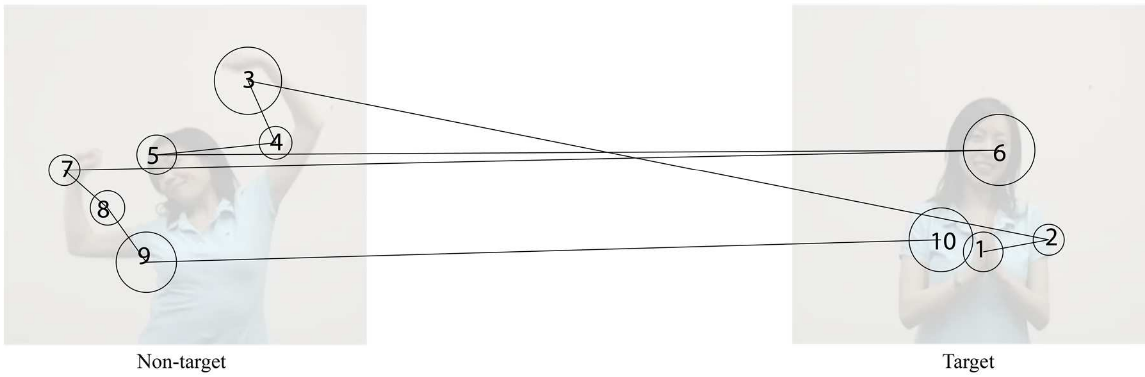


Figure 1: This shows a schematic view of gaze patterns. Each circle represents a fixation, with the diameter of the circle corresponding to the duration of that fixation. The number in the center of the circle represents the order in which the fixations occurred. Fixation density is the ratio of the number of fixations on the non-target (6) to number of fixations on the target (4), giving us a fixation density of 1.5 in this example. The second exploratory measure is the duration of the first fixation on the non-target, which would be indicated above by the diameter of the circle labeled 3.

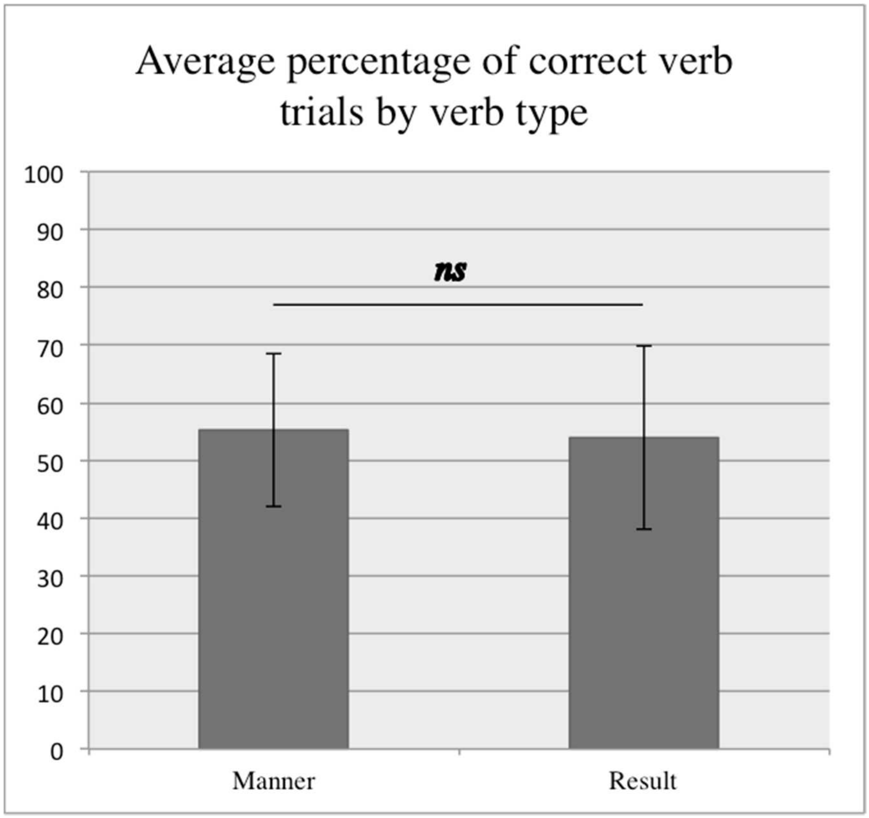


Figure 2: Average percentage of correct verb trials by verb type.

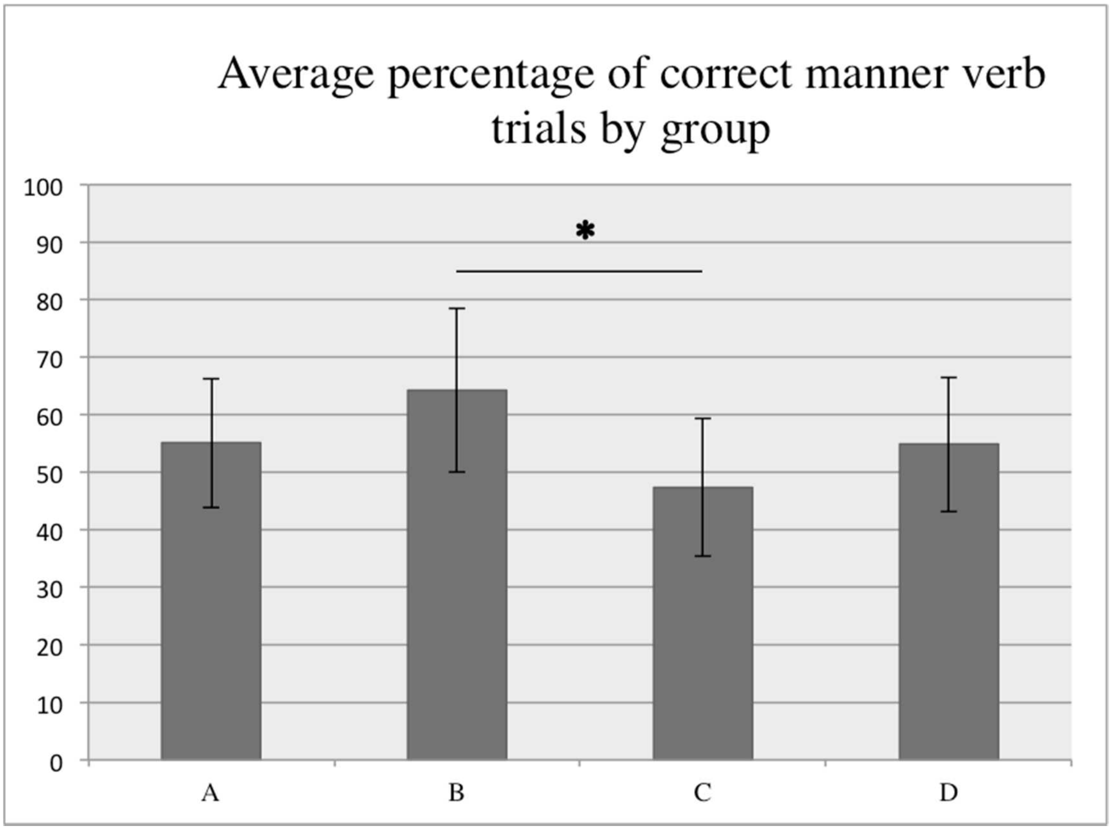
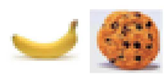


Figure 3: Average percentage of correct manner verb trials by group, with a significant difference between groups B and C (p=.021).

APPENDIX 1: Schematic of one trial

Image on Screen						
Time on screen	2 seconds	2 seconds	.25 seconds	3 seconds	4 seconds	6 seconds
Auditory input				Do you see?	Can you show me the cookie?	Where is the cookie? Let's find the cookie!
Subphase Name	Cookie-Banana-Salience-R	Cookie-Banana-Salience-L	Cookie-Banana-White	Cookie-Banana-Salience-Both	Cookie-Banana-Star	Cookie-Banana-Response

APPENDIX 2.A: Counterbalancing for Groups A and B

A				
Trial Type	Trial Name	Target	Side Correct	Side First
Attn	Duck			
N	Cookie-Banana	Cookie	L	R
N	Goldfish-Donut	Donut	R	R
N	Firetruck-Bird	Firetruck	L	L
V	Feed-Hug	Hug	R	L
V	Pour-Drink	Pour	L	L
V	Wash-Rock	Wash	L	R
Attn	Rattle			
V	Cut-Tie	Tie	R	L
N	Crab-Pancakes	Crab	L	R
V	Eat-Push	Eat	L	L
V	Run-Jump	Jump	R	R
V	Shake-Open	Open	R	L
V	Read-Rip	Read	L	R
Attn	Baby			
N	Rocketship-Giraffe	Giraffe	R	R
V	Stretch-Clap	Clap	R	R
V	Roll-Bounce	Roll	L	L
V	Lift-Pull	Lift	L	R
V	March-Spin	Spin	R	L
N	Squirrel-Grapes	Grapes	R	R
Attn	Elephant			
V	Dance-Cry	Dance	L	L
V	Drop-Bite	Bite	R	L
V	Kiss-Tickle	Tickle	R	R
V	Squeeze-Blow	Squeeze	L	R
N	Orange-Airplane	Orange	L	L
V	Kick-Throw	Throw	R	L
V	Lick-Break	Lick	L	R

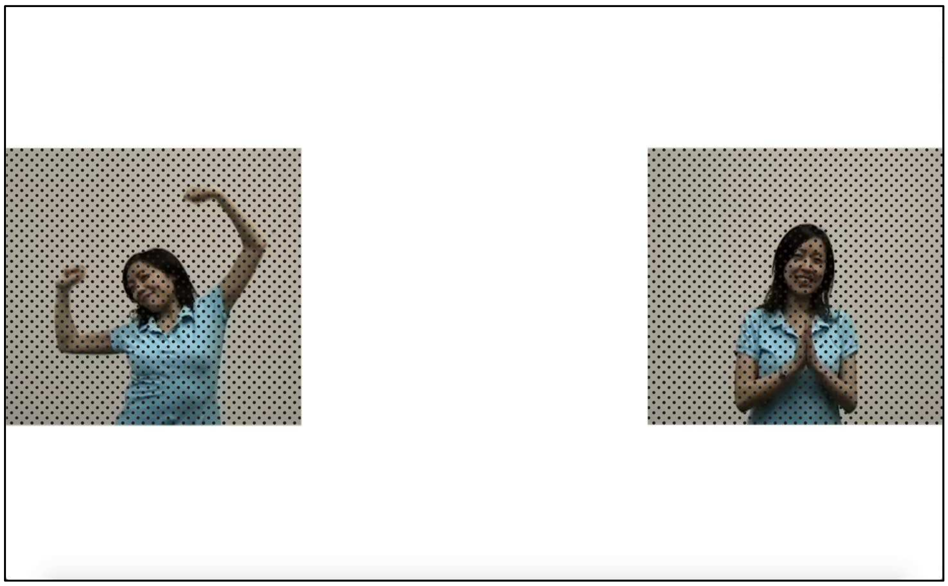
B				
Trial Type	Trial Name	Target	Side Correct	Side First
Attn	Duck			
N	Cookie-Banana	Banana	R	R
N	Goldfish-Donut	Goldfish	L	R
N	Firetruck-Bird	Bird	R	L
V	Feed-Hug	Feed	L	L
V	Pour-Drink	Drink	R	L
V	Wash-Rock	Rock	R	L
Attn	Rattle			
V	Cut-Tie	Cut	L	L
N	Crab-Pancakes	Pancakes	R	R
V	Eat-Push	Push	R	L
V	Run-Jump	Run	L	R
V	Shake-Open	Shake	L	L
V	Read-Rip	Rip	R	R
Attn	Baby			
N	Rocketship-Giraffe	Rocketship	L	R
V	Stretch-Clap	Stretch	L	R
V	Roll-Bounce	Bounce	R	L
V	Lift-Pull	Pull	R	R
V	March-Spin	March	L	L
N	Squirrel-Grapes	Squirrel	L	R
Attn	Elephant			
V	Dance-Cry	Cry	R	L
V	Drop-Bite	Drop	L	R
V	Kiss-Tickle	Kiss	L	L
V	Squeeze-Blow	Blow	R	R
N	Orange-Airplane	Airplane	R	L
V	Kick-Throw	Kick	L	R
V	Lick-Break	Break	R	R

APPENDIX 2.B: Counterbalancing for Groups C and D

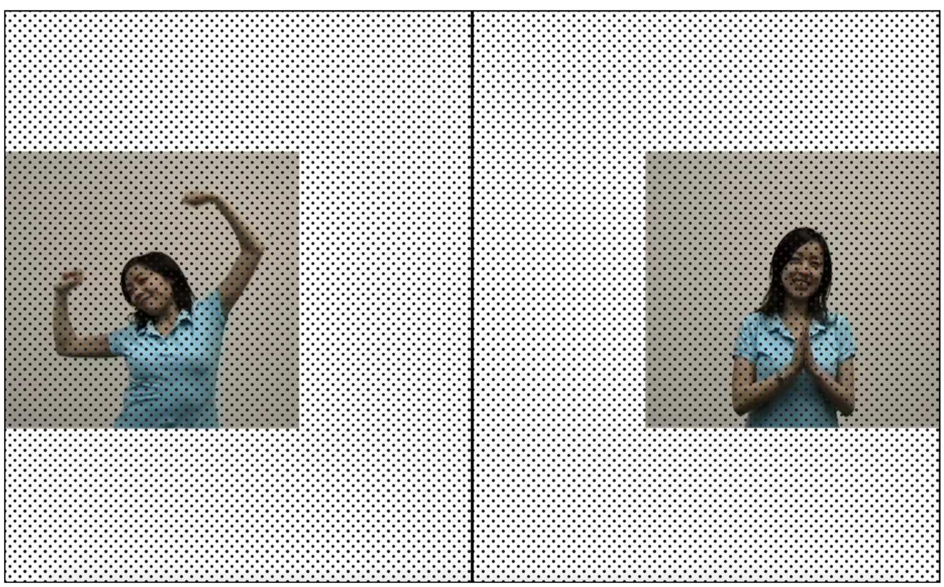
C				
Trial Type	Trial Name	Target	Side Correct	Side First
Attn.	Duck			
N	Orange-Airplane	Airplane	R	L
N	Squirrel-Grapes	Squirrel	L	R
N	Rocketship-Giraffe	Rocketship	L	R
V	Lick-Break	Break	R	R
V	Kick-Throw	Kick	L	R
V	Squeeze-Blow	Blow	R	R
Attn	Rattle			
V	Kiss-Tickle	Kiss	L	L
N	Crab-Pancakes	Pancakes	R	R
V	Drop-Bite	Drop	L	R
V	Dance-Cry	Cry	R	L
V	March-Spin	March	L	L
V	Lift-Pull	Pull	R	R
Attn	Baby			
N	Firetruck-Bird	Bird	R	L
V	Roll-Bounce	Bounce	R	L
V	Stretch-Clap	Stretch	L	R
V	Read-Rip	Rip	R	R
V	Shake-Open	Shake	L	L
N	Donut-Goldfish	Goldfish	L	R
Attn	Elephant			
V	Run-Jump	Run	L	R
V	Eat-Push	Push	R	L
V	Cut-Tie	Cut	L	L
V	Wash-Rock	Rock	R	L
N	Banana-Cookie	Banana	R	R
V	Pour-Drink	Drink	R	L
V	Feed-Hug	Feed	L	L

D				
Trial Type	Trial Name	Target	Side Correct	Side First
Attn	Duck			
N	Orange-Airplane	Orange	L	L
N	Squirrel-Grapes	Grapes	R	R
N	Rocketship-Giraffe	Giraffe	R	R
V	Lick-Break	Lick	L	R
V	Kick-Throw	Throw	R	L
V	Squeeze-Blow	Squeeze	L	R
Attn	Rattle			
V	Kiss-Tickle	Tickle	R	R
N	Crab-Pancakes	Crab	L	R
V	Drop-Bite	Bite	R	L
V	Dance-Cry	Dance	L	L
V	March-Spin	Spin	R	L
V	Lift-Pull	Lift	L	R
Attn	Baby			
N	Firetruck-Bird	Firetruck	L	L
V	Roll-Bounce	Roll	L	L
V	Stretch-Clap	Clap	R	R
V	Read-Rip	Read	L	R
V	Shake-Open	Open	R	L
N	Goldfish-Donut	Donut	R	R
Attn	Elephant			
V	Run-Jump	Jump	R	R
V	Eat-Push	Eat	L	L
V	Cut-Tie	Tie	R	L
V	Wash-Rock	Wash	L	R
N	Cookie-Banana	Cookie	L	R
V	Pour-Drink	Pour	L	L
V	Feed-Hug	Hug	R	L

APPENDIX 3: AOI coding, narrow versus broad



Narrow coding – dots represent what would count as a look to the target or non-target



Broad coding – dots represent what would count as a look to the target or non-target, with a midline split for left and right

REFERENCES

Behrend, D. A. (1990). The development of verb concepts: Children's use of verbs to label familiar and novel events. *Child Development* 61, 681–696.

Brady, N. C., Anderson, C. J., Hahn, L. J., Obermeier, S. M., & Kapa, L. L. (2014). Eye tracking as a measure of receptive vocabulary in children with autism spectrum disorders. *Augmentative and Alternative Communication*, *30*(2), 147–159. doi:10.3109/07434618.2014.904923

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic Theory and Psychological Reality* (pp. 264–293). Cambridge, MA: The MIT Press.

Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology*, *38*(6), 967–978. doi:10.1037/0012-1649.38.6.967

Cocking, R. R., & McHale, S. (1981). A comparative study of the use of pictures and objects in assessing children's receptive and productive language. *Journal of Child Language*, *8*, 1–13.

Dunn, L. M., Dunn, D. M., & Pearson Assessments. (2007). PPVT-4: Peabody picture vocabulary test. Minneapolis, MN: Pearson Assessments.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., … Stiles, J. (1994). Variability in Early Communicative Development. *Monographs of the Society for Research in Child Development*, *59*(5), 174–185.

Fenson, L., & Paul H. Brookes Publishing Co. (2007). MacArthur-Bates communicative development inventories. Baltimore, MD: Paul H. Brookes Pub. Co.

Fernald, A., Perfors, A., & Marchman, V. a. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*(1), 98–116. doi:10.1037/0012-1649.42.1.98

Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, *9*(3), 228–231.

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. M. Fernandez, & H. Clahsen (Eds.), *Developmental Psycholinguistics* (pp. 97–135). Philadelphia: John Benjamins Publishing Company.

Gentner, D. (1978). On relational meaning: The acquisition of verb meaning. *Child Development*, *49*(4), 988–998. doi:10.2307/1128738

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, *73*(2), 135–176. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10580161

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, *14*, 23–45. doi:10.1017/S030500090001271X

Gropen, J., Pinker, S., Hollander, M., & Goldberg, R. (1991). Affectedness and direct objects: The role of lexical semantics in the acquisition of verb argument structure. *Cognition, 41*, 153–195.

Gupta, P., & Tisdale, J. (2009). Word learning, phonological short-term memory, phonotactic probability and long-term memory: towards an integrated framework. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1536), 3755–3771. doi:10.1098/rstb.2009.0132

Henderson, J. M., Weeks, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 210–228.

Horst, J. S., & Samuelson, L. K. (2008). Fast Mapping but Poor Retention by 24-Month-Old Infants, *13*(2), 128–157. doi:10.1080/1525O(xH)701795598

Horvath, S., Rescorla, L., & Arunachalam, S. (2015). The expressive verb vocabulary of typically developing versus language delayed two-year olds. *Poster Presented at Boston University Scholar's Day*.

Houston-Price, C., Mather, E., & Sakkalou, E. (2007). Discrepancy between parental reports of infants' receptive vocabulary and infants' behaviour in a preferential looking task. *Journal of Child Language*, *34*, 701–724. doi:10.1017/S0305000907008124

Imai, M., Li, L., Haryu, E., Okada, H., Hirsh-pasek, K., Golinkoff, R. M., & Shigematsu, J. (2008). Novel Noun and Verb Learning in Chinese-, English-, Japanese-speaking children. *Child Development*, *79*(4), 979–1000.

Lee, J. (2011). Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics*, *32*, 69–92. doi:10.1017/S0142716410000299

Marchman, V. A., & Bates, E. A. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language*, *21*, 339–366.

Markman, E. M. (1991). The whole-object, taxonomic and mutual exclusivity assumptions as initial constraints on word meanings. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought*.

Munro, N., Baker, E., McGregor, K., Docking, K., & Arculi, J. (2012). Why Word Learning is not Fast. *Frontiers in Psychology*, *3*(February), 41. doi:10.3389/fpsyg.2012.00041

Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.

Salvucci, D. D., Salvucci, D. D., Goldberg, J. H., & Goldberg, J. H. (2000). Identifying Fixations and Saccades in Eye-Tracking Protocols. *Proceedings of the Eye Tracking Research and Applications Symposium*, 71–78. doi:10.1145/355017.355028

Stickgold, R., & Walker, M. P. (2005). Memory consolidation and reconsolidation: what is the role of sleep? *Trends in Neurosciences*, *28*(8), 408–415. doi:10.1016/j.tins.2005.06.004

Styles, S., & Plunkett, K. (2009). What is "word understanding" for the parent of a one-year-old? Matching the difficulty of a lexical comprehension task to parental CDI report. *Journal of Child Language*, *36*, 895–908. doi:10.1017/S0305000908009264

Swensen, L. D., Kelley, E., Fein, D., & Naigles, L. R. (2007). Processes of language acquisition in children with autism: Evidence from preferential looking. *Child Development*, *78*(2), 542–557. doi:10.1111/j.1467-8624.2007.01022.x

Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology*, *32*(3), 492–504. doi:10.1037//0012-1649.32.3.492

Tardif, T., Gelman, S. A., & Xu, F. (1999). Putting the " Noun Bias " in Context : A Comparison of English and Mandarin Measuring Early Vocabularies : Observational, *70*(3), 620–635.

Venker, C. E., Kover, S. T., & Weismer, S. E. (2015). Fast Mapping Predicts Differences in Concurrent and Later Language Abilities Among Children with ASD. *Journal of Autism and Developmental Disorders*, *46*(3), 1118–1123. doi:10.1007/s10803-015-2644-x

Werchan, D. M., & Gómez, R. L. (2014). Wakefulness (not sleep) promotes generalization of word learning in 2.5-year-old children. *Child Development*, *85*(2), 429–436. doi:10.1111/cdev.12149

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics: Research article. *Psychological Science*, *18*(5), 414–420. doi:10.1111/j.1467-9280.2007.01915.x

CURRICULUM VITAE

Matthew James Valleau (b. 1988)

Education

Northeastern University, Bachelor of Arts (*summa cum laude*), 2011

Major: Spanish and French. *Minor*: Linguistics

Boston University, Master of Arts, 2014

Major: Applied Linguistics

Boston University, Master of Science, 2016

Major: Speech-Language Pathology

Presentations

Matthew Valleau and Chelsea DeSantis. Japanese and American Reactions to Socially-Marked Variation in English, McGill's Canadian Conference for Linguistics Undergraduates, Montréal, Canada

Matthew Valleau, Michael Frank and Sudha Arunachalam. Cross-Linguistic differences in maternal supports for word learning. Poster presented at the International Conference on Infant Studies, Berlin, Germany.

Publications

William Orman & Matthew James Valleau (Eds.). 2014. *Proceedings of the 38th Annual Boston University Conference on Language Development.* Boston, MA: Cascadilla Press

Please send all correspondence to:

Matthew James Valleau

valleaum@bu.edu

Boston University

Department of Speech, Language and Hearing Sciences

635 Commonwealth Avenue

Boston, MA 02215