5-9-2015

Effects of Auditory and Visual Variability on Word Learning in Children: A Pilot Study

Kelly A. Casey
University of Connecticut - Storrs, kelly.a.casey@uconn.edu

Effects of Auditory and Visual Variability on Word Learning in Children: A Pilot Study

Kelly Ann Casey

B.A., University of Connecticut, 2013

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Arts

At the

University of Connecticut

2015

APPROVAL PAGE

Master of Arts Thesis

Effects of Auditory and Visual Variability on Word Learning in Children: A Pilot Study

Presented by

Kelly Ann Casey

Major Advisor ______________________________
Rachel M. Theodore, Ph.D.

Associate Advisor ______________________________
Bernard G. Grela, Ph.D.

Associate Advisor ______________________________
Tammie J. Spaulding, Ph.D.

University of Connecticut

2015

Acknowledgments

I would like to thank Dr. Rachel M. Theodore for her support as my thesis advisor throughout this process. She provided unwavering assistance and guidance throughout the entire experience. Her passion and dedication for research has fostered my appreciation for further investigation in the field of speech and language production and perception.

I would also like to thank Dr. Kristin A. Vasil-Dilaj for her contributions to this project. She provided unique knowledge and insight from the area of auditory perception which strengthened our research design and investigation.

I would like to thank fellow graduate students Shayna Marmon and Alexandra Bohner for all of their help in the preparation and running of this experiment, as well as their assistance in data collection and analysis. I would also like to acknowledge the assistance of the undergraduate research assistants, Taylor Leach, Alexandra Gagas, Kylie Hill, Marykate Bisaillon, Devin Roscillo, Emily Butterworth, and Elana Katz, for their hard work and their help in acquiring stimulus items and running this experiment. This study could not have been completed without their dedication and contribute ons. I would also like to acknowledge the input and unique perspectives provided by my reading committee, Dr. Bernard G. Grela and Dr. Tammie J. Spaulding.

Lastly, I would like to thank my family and friends for their constant encouragement and support throughout my experience in this program for the past two years. Their confidence in me and their advice has provided me with motivation to reach both personal and academic goals.

This research was funded by a Clinical Research Grant from the University of Connecticut Speech and Hearing Clinic to Dr. Rachel M. Theodore and Dr. Kristin A. Vasil-Dilaj and by start-up funds to Dr. Rachel M. Theodore from the College of Liberal Arts and Sciences at the University of Connecticut.

Table of Contents

Abstract

For infants, acquiring vocabulary for nouns is a dynamic, complex process that involves pairing an auditory token with a visual referent. This process is computationally complex because the acoustic information produced for a verbal production of any given noun varies considerably due to factors including the person who is speaking, speaking rate, and linguistic context. Likewise, visual referents are also variable in characteristics such as size, shape, material, and color. Research suggests that variability in either the auditory or visual domains can facilitate early word learning. However, the role of simultaneous variability in these domains on noun learning remains unexplored. Using a 9-week training study, we examined the effects of auditory and visual variability on word learning and generalization in 12 children ages 16- to 23-months in order to collect pilot data for a larger-scale investigation. All children were taught 12 nouns and were randomly assigned to one of four training conditions: low visual and low auditory variability, low visual and high auditory variability, high visual and low auditory variability, or high visual and high auditory variability. High versus low auditory variability was manipulated by presenting ten talkers versus one talker, respectively. High versus low visual variability was manipulated by presenting variable, dissimilar exemplars versus highly similar exemplars, respectively. The results to date suggest that high levels of variability in the visual domain facilitated learning of trained items but did not influence the ability to generalize that category to novel visual exemplars. Moreover, overall vocabulary development appeared to be facilitated by high variability in the auditory domain. These findings provide promising pilot data for understanding how visual and auditory variability influence word learning not only in the laboratory, but also in the real-world linguistic environment.

Note

This thesis reflects a working project in collaboration with Dr. Rachel M. Theodore, Dr. Kristin A. Vasil-Dilaj, and Ms. Shayna Marmon. The manuscript related to this thesis project will be submitted following final data collection and authorship will be shared by those named above.

Introduction

Language acquisition is a dynamic, complex process in which children gradually acquire a wide variety of language skills across various components such as, phonology, morphology, syntax, semantics, and pragmatic use of language. Regarding these language components, specifically semantics, one major achievement in language acquisition is acquiring vocabulary for a wide variety of nouns. The vocabulary acquisition process for nouns involves a systematic pairing of an auditory stimulus with a visual referent. For example, when learning the word "dog," the child must recognize that a specific acoustic token refers to an object in their environment, or a specific visual referent.

However, a theoretical understanding of this pairing process is not straightforward because there is variability in both the auditory and visual domains for a given noun. It is rare for a typically developing child to be exposed to auditory input from one speaker only and visual input of only one specific type of object. It is more likely that children will hear the word "dog" spoken by various talkers and will likely be exposed to a variety of visual referents that all fall within the category of "dog." Accordingly, models of language acquisition must describe the mechanisms that support word learning given the variability in mapping between a visual object and its auditory referent. The goal of the current work is to contribute to such an account.

Sources of inherent auditory complexity

Variability across aspects of the acoustic token for a given word makes this word learning process more complex. Two speech phenomena have been suggested as requirements for infants to perceive invariant relation between different acoustic tokens: categorical perception and the capacity to recognize the same utterances produced by different speakers (Jusczyk, 1986). The

child must realize that despite considerable variability in an acoustic token for a word, its meaning remains consistent. The acoustic token can vary due to several factors, such as speaking rate (Miller, 1981), phonetic context (Delattre, Liberman, & Cooper, 1955), and who is speaking (Jusczyk, 1986; Klatt, 1986).

Variability in speech rate can alter articulation times for vowels and consonants which, in turn, may result in changes in the acoustic and perceptual properties of sounds (Miller, 1981). A study conducted by Lindblom (1963) examined the effect of changes in speech tempo and stress on eight Swedish vowels. Results demonstrated that as the rate of speech during a given word was increased (and thus the duration of the syllable decreased), there was a tendency for the formant frequency values of the vowel to undershoot their target, resulting in a vowel reduction that produced a more "schwa-like" production. However, another study conducted by Verbrugge and Shankweiler (1977) indicated that while increased speech rate similarly resulted in decreased average syllable duration, the formant frequency values did not demonstrate a significant trend towards a reduced vowel space. While vowel reduction does not always occur in the face of increased speech rate, it is crucial to consider the implications for vowel perception when it does indeed occur. Because the formant frequencies for a given vowel may not be achieved due to increased rate, the listener is required to account for this increased rate of speech and compensate for these changes to specify vowel identity. Because vowels contribute greatly to determine the perception and meaning of words, changes as a result of speech rate add to the complexity of this task. Speaking rate also can result in modification in the acoustic properties of consonant distinction. In a study conducted by Summerfield (1975a), the effect of rate of speech on the production and perception of initial position voiced and voiceless consonants was examined. Results from this study indicated that variations between slow, normal, and fast rate of speech

yielded different voice-onset-times (VOTs) depending on whether or not the consonant was voiced or voiceless in the initial position of the syllable. While there was rarely an interruption noted in voicing for a voiced consonants /b/ and /g/, there was a significant decrease in VOT of voiceless stops when rate of the utterance was increased. Additionally, the difference between average VOT values of voiced and voiceless stops diminished when the rate of speech increased. These changes in VOT as a result of variation in speech rate can make the process of sound discrimination within a word more complex because listeners are required to adjust their decision criteria for sounds (or words) in line with systematic variability when exposed to varied speaking rates (Pols, 1986).

In addition to variation in speech rate, the process of learning words may be influenced by variation in phonetic context. In spoken language, the way that sounds are discriminated and perceived can be broken into various levels: phonemic variation, the position of the sound in the word, the local context, and the remote context. Phonemic variation refers to acoustic differences for the same phoneme. Local context refers to variation in sound perception at the syllabic level perhaps due to coarticulation or assimilation. Remote context refers to variability at the word and sentence levels influenced by prosodic aspects, such as tempo, stress, and accent (Pols, 1986). Delattre, Liberman, and Cooper (1955) discuss findings of their previous work which indicate that transitions, or frequency shifts, of the second formant of a vowel serves as cues to identify certain voiced stop consonants, such as /b/, /d/, /g/, voiceless stop consonants, such as /p/, /t, /k/, and nasal consonants /m/, /n/, and /ŋ/. Additionally, they found that first formant loci of the vowel correspond with manner of articulation and second formant loci generally aid in identification of place of articulation. These findings contribute to knowledge of speech and sound perception in regards to context, and more specifically, how transitions in vowel

frequency influence our identification of the preceding consonant.

Yet another source of variability in speech production concerns individual talker differences in phonetic properties of speech. Subsequently, another aspect of auditory variability to consider during vocabulary acquisition is the differences in speech production between speakers. While it is important to discriminate between speech segments in a variety of contexts, it is equally important to demonstrate the ability to recognize the same segment when spoken by different speakers (Jusczyk, 1986). This becomes especially important when considering the effect auditory input has on verbal output for infants. An infant's ability to recognize separate productions as phonetically equivalent when produced by different talkers is an indicator of their capacity for vocal imitation (Kuhl, 1983). Acoustic differences between speakers can be attributed to the following factors: vocal tract length and shape, sex, articulatory habits, and dialectical differences (Klatt, 1986). The variability in vocal tract length across speakers results in changes in the distribution of sound energy and frequency, which results in the same word or sound being produced differently. However, while the listener may detect these differences, they are still expected to recognize that these are productions of the same word or sound. Kuhl's (1983) findings suggest that children as early as six months were able to appropriately discriminate between vowels /a/ and /□/ and maintain constancy for this vowel discrimination across talkers, including males, females, and children, as well as rising and falling pitch contours. Thus, infants very early on demonstrate the ability to identify events as discriminately different but linguistically equivalent across various talkers and ages.

Sources of inherent visual complexity

Along with aspects of auditory complexity, inherent variability in the visual domain adds

to the complex process of word learning. When a child is pairing a given auditory stimulus with a visual stimulus, they are systematically attending to various properties of that object (Soja, Carey, & Spelke, 1991). For example, when a child hears the word “dog,” the child must decide if the word refers to the animal as a whole, a part of the animal, an action performed by the animal, or a property or characteristic of the animal, among many other options. When children are learning novel nouns, the visual referent that corresponds to the given noun may vary in shape, material, size, and color. The child is expected to realize that although there are differences among visual exemplars for one given word or category, they still fall within that category. Previous work suggests that upon seeing an object for the first time and hearing it named, such as a *tractor*, it is likely that the child will recognize other tractors and extend this category to novel tractors even if the items differ from the original tractor in size or color (Clark, 1973; Mervis, 1987). Previous work has also indicated that for non-solid objects, children as young as two years of age exhibit the ability to generalize a name based on color and texture (Soja, et al., 1991).

During this process of extending a name to a novel object, one physical attribute that children attend to is the shape of the object being named. Previous work has suggested that while children between the ages of seventeen to thirty-three months are beginning to understand how a noun is linked with an object category, they tend to extend names to novel, solid objects that are similar in shape (Gershkoff-Stowe & Smith, 2004; Samuelson & Smith, 1999; Smith, 2002; Soja et al., 1991). One way in which researchers have examined this tendency for children to extend words to novel objects on the basis of shape is by using a novel word learning or generalization task. The children were presented with a visual exemplar and were provided the name corresponding to that object (e.g. “This is a dax.”). They were then shown other exemplars that

shared a similar shape and were asked to pick or point to the object of the word that was spoken, which would indicate whether or not they were generalizing the previously learned word to other objects of the same shape. Solid objects are generally considered to be concrete, individual objects that are bound by the amount of space they occupy, whereas non-solid objects are ones that are continuous and non-discrete (Soja et al., 1991). These studies provide support for the notion that when children are presented with several solid objects that vary among attributes such as color, texture, and shape, they tend to generalize the word they previously learned for the noun to objects that were similar in shape. Additionally, some children demonstrated an acceleration of noun production outside of the lab environment, suggesting that there is a relationship between the rate of noun acquisition and attention to shape (Gershkoff-Stowe & Smith, 2004).

While previous studies provide support for a shape bias when learning names for nouns, other physical attributes like material and color influence the word and categorization process as well. Specifically, when the object being learned was made of non-solid materials, such as glue or gel, children generally no longer categorize the object by shape but instead tend to categorize the object based on its color or material or substance (Soja, et al., 1991). Children as young as two years of age demonstrated the ability to ignore the shape bias seen for solid objects and instead generalized exemplar name based on color or material. This suggests that during the process of generalizing previously learned words to novel objects, young children can shift their attention between various physical properties to categorize the object.

Typical word learning development

Despite these complexities found in the variability in auditory and visual domains,

typically developing children acquire new vocabulary with ease. Between the ages of eighteen and twenty-four months, there is generally a rapid expansion in expressive naming abilities in children, with words consisting primarily of nouns (Dromi, 1986; Nelson, 1973; Samuelson & Smith, 1999). Dromi (1986) suggested that when determining how meanings are conveyed by words and in which contexts they would be used appropriately, children at this stage identify the following: words convey consistent meanings, words differ based on whether they are used for nouns, actions, or relations, words should correspond with adults' productions in form and meaning, words are often associated with more than one specific referent, and the meaning of two words contrast with one another. Additionally, regarding vocabulary comprehension, previous work has indicated that after receiving just nine exposures to a new object label within a five minute training session, both thirteen and eighteen month old children demonstrated comprehension of new words (Woodward, Markman, and Fitzsimmons, 1994). This suggests that children can learn a new object label fairly quickly before they enter the aforementioned expressive vocabulary explosion. Thus, even in the face of inherent complexities in both auditory and visual domains, children demonstrate the ability to both comprehend and produce new words with remarkable ease.

Word learning and auditory variability

Previous work has suggested that increased variability in both the auditory and visual domains may separately facilitate the word learning process in children. Previous work has indicated that infants demonstrate difficulty learning lexical neighbors, or words that are considered phonologically similar but contrast by a single phoneme, such as *bih* versus *dih* (Stager & Werker, 1997). In a series of experiments, children fourteen months of age were

habituated to word-object matches for words that varied by a single phoneme (*bih* and *dih*) as well as by several phonemes (*lif* and *neem*). Following habituation, the subjects were presented with same trials, where objects matched the word they learned in habituation, or switch trials, where the object did not match the word it was habituated with. The subjects could determine when there was a mismatch or switch when the word pair differed by multiple phonemes but consistently failed to determine a mismatch when the words varied by a single phoneme. More recent work has determined that learning of these lexical neighbors can be facilitated when the child is exposed to several talkers producing the contrast as opposed to just hearing the contrast from a single speaker (Rost & McMurray, 2009). Two experiments were conducted with children fourteen months of age: one in which the subjects were habituated to auditory stimuli consisting of /buk/ versus /puk/ produced by a single speaker and another where children were habituated and tested on the same auditory stimuli produced by eighteen talkers. Results of the first experiment were consistent with the findings of Stager and Werker (1997), in that the infants were unable to detect differences between the lexical neighbors. Results of the second experiment, however, indicated that the subjects succeeded at learning two phonologically similar words when presented with multiple exemplars in the switch task. Therefore, variability in the auditory domain in this case facilitated learning of more robust lexical categories.

Given that multi-talker auditory input has been shown to aid in minimal pair learning, it is important to next consider which components of talker variability contributed to this learning. Previous work has demonstrated that infants as young as seven and a half months of age could generalize recognition of familiarized words when there were changes in amplitude but not pitch. However, by nine months of age, infants demonstrated this skill in the face of variation in both amplitude and pitch (Singh, White, & Morgan, 2008). A study conducted by Rost and McMurray

(2010) investigated both phonologically contrastive aspects of speech, such as voice-onset-time (VOT), and non-contrastive components of speech, such as pitch, prosody, and voice quality through three experiments. Experiment 1 was similar to the previous switch tasks with the exception that the VOTs of the /buk/ and /puk/ auditory exemplars, produced by a single speaker, were manipulated along a continuum. Experiment 2 consisted of a similar task with the exception that VOT, burst amplitude, and fundamental frequency were manipulated. These two experiments, therefore, varied these phonologically contrastive components and controlled for variation among non-contrastive components. In Experiment 3, the auditory exemplars from the eighteen speakers used in the Rost and McMurray (2009) study were used but the VOTs were manipulated to all be 2 ms to control for variation in contrastive components. Variation across prosody and fundamental frequency was examined. Results from Experiment 1 and 2 revealed that variation along contrastive cues alone did not facilitate word learning of the minimal pairs, whereas results from Experiment 3 indicated that variability among non-contrastive speech components facilitated minimal pair learning in the switch task. Therefore, these aspects of auditory variability between speakers facilitated learning of these two words.

Additional work has contributed to support for variability in the auditory domain facilitating word recognition. Singh (2008) examined the effects of both low and high auditory variability during four word recognition experiments. The source of variability across these experiments was vocal affect. In Experiment 1, subjects were familiarized with one word spoken with variable affect (variable) and another word with a single emotion (constant). The words were dispersed in passages where the affect always mismatched the affect of the constant word. The infants were able to recognize words that were familiarized in the variable affect, lending support to the notion that increased variability in the auditory domain assists word recognition.

Word learning and visual variability

Conflicting viewpoints exist regarding whether less variability or more variability in the visual domain aids in establishing categories for objects. Oakes, Coppage, and Dingle (1997) examined the effects of perceptual similarity in object categorization for children ten and thirteen months of age. Over the course of three experiments, children aged ten months were better able to make category distinction between land and sea animals when the exemplars within each category were similar, rather than variable. However, children aged thirteen months were able to make distinctions between groups with lower levels of perceptual similarity, or in other words, more variable exemplars within each category. Based on these findings, it is possible that as infants age, they become less reliant on perceptual similarity to form exclusive categories and eventually can benefit from exemplar variability.

More recent work suggests variability in the visual domain has, in fact, been shown to facilitate language acquisition in young children. Researchers examined the influence of different levels of similarity across related events, or cross-situational information, on learning of novel verbs in Korean and English-speaking children (Childers & Paik, 2009). Children aged two and half to three and a half were familiarized with four new words that were associated with four different target events. The familiarization task included the experimenter enacting an event while producing three sentences with the novel verb, *blick* in this example: (1) “Look, I’m going to *blick* it,” (2) “I’m *blicking* it,” and (3) “I *blicked* it.” The children were then given a chance to enact the event and say the verb. During the test phase, children were assigned to one of two conditions: one in which they saw three new events and objects that were similar to the target event (close comparison group) or the other condition in which they saw three new events and objects that were dissimilar from the target event but achieved the same result as the target event

(far comparison group). Children again were asked to reenact the actions and use the verb in response to questions regarding what they were doing. Results demonstrated that children who experienced varied events enacted the event using more varied objects than those who experience similar events. Therefore, they were generalizing the verb in a way that was appropriate when using a variety of different objects. Additionally, children who experienced similar events produced less extensions of the verb and extended the new word less frequently in general.

Variability in visual exemplars has not only been seen to facilitate learning of verbs but also for names of concrete noun categories. A longitudinal study, which consisted of nine weekly sessions and a one-month follow-up, was conducted to examine the role of visual exemplar variability on word acquisition and generalization in sixteen eighteen-month-old children (Perry et al., 2010). The children were taught words corresponding to twelve categories and were assigned to one of two conditions: half of the children were assigned to a condition where they were taught the categories with sets of highly similar exemplars (tight condition) and the other half were assigned to a condition where they were taught with variable, dissimilar visual exemplars (variable condition). They were taught the names of the twelve categories through naturalistic play, where they were shown either three highly similar or three variable exemplars of each category depending on their condition. The children were then tested in four different areas: (a) learning of exemplar names they were trained on, (b) generalization of learned labels to novel exemplars, (c) novel noun generalization, and (d) overall vocabulary acquisition. Results indicated that children in the tight condition learned the individual exemplars equally as well as children in the variable condition but children in the variable condition were more likely to generalize the names to novel stimuli. In this case, high levels of variability supported an

increased ability to generalize what has been learned to new exemplars. Additionally, children in the variable condition acquired new words at a significantly faster rate between Week 9 and the 1-month follow-up, resulting in a greater productive vocabulary by the end of the study. This finding demonstrates a significant effect of learning in laboratory training extending to learning in environments outside of the lab.

Other work has examined the effects of word learning of adjectives, specifically color words. In a study conducted by Thom and Sandhofer (2009), twenty month old children were trained and tested on color word based on their assignment to one of three conditions: a two-word condition, a four-word condition, and a six-word condition. The study consisted of eleven sessions. In session one, a pretest was conducted for color comprehension and production. The next eight sessions focused on training of the words, in which the experimenter showed two objects of the same color and repeated the color name ten times in total, as well as introducing a third object that differed in color. Session ten consisted of a post-test identical to that of the comprehension pretest except the trained words were tested twice. Session eleven was an extension test to determine if the child would map a novel, untrained color word to its referent in a single naming session. Results indicated that the children who were trained in more color words were better able to extend new words to novel instances than children who were trained in fewer words. In this case, more variability in the colors learned facilitated generalization in new instances. Researchers considered the possibility that learning the labels of multiple exemplars within a category may promote attention to relevant features that determine whether or not the object falls within that category.

When considering the amount of variation that is inherently present in the auditory and visual domains while children are learning words, they appear to be quick and skilled learners.

The previous work has suggested that hearing phonologically similar words produced by several different talkers as opposed to just one talker facilitates the ability to develop lexical neighbors Rost & McMurray, 2009). Additionally, variation in acoustic information found between talkers, such as pitch, voice quality, prosody, and affect has been shown to facilitate word learning and recognition (Rost & McMurray, 2010; Singh, 2008). Variability in the visual domain has been shown to facilitate the learning of categories in infants and novel verbs in toddlers. It has also been shown to facilitate young children's ability to generalize previously learned nouns to novel visual exemplars and fostered vocabulary growth inside and outside of the laboratory (Perry et al., 2010). While research has investigated the independent contributions of variability each of these domains, the effects of simultaneous variability has not yet been examined. Our study focuses on the influence of simultaneous variability in both the auditory and visual domains and how this affects word learning and generalization of novel nouns.

Current study

The goal of the current study was to examine vocabulary acquisition for nouns in typically-developing children between 16 and 23 months of age. All subjects participated in a nine-week word learning study in which each subject was randomly assigned to one of four experimental conditions: low auditory and low visual variability, low auditory and high visual variability, high auditory and low visual variability, and high auditory and low visual variability. While previous work has examined the independent contributions of auditory and visual variability in vocabulary acquisition, these sources of variability have not been examined simultaneously. Thus, our study aims to address the following research questions: (1) does variability in the auditory domain facilitates word learning as has been demonstrated for visual

variability, and (2) does simultaneous variability in both the auditory and visual domains promote enhanced word learning compared to variability in either domain alone?

Methods

Participants

The participants in this study were twelve typically developing children ranging from 16 to 23 months of age. All children were recruited from monolingual English-speaking homes and had no history of speech, language, or hearing disorders as confirmed by parent report. Additionally, hearing screenings were administered for each child at least three times throughout the duration of the study in order to confirm parent report and that there was no presence of short-term hearing loss due to illness. All participants were recruited via flyers, a weekly university newsletter, and a research participant contact database. The children were randomly assigned to one of four conditions formed by crossing the two independent variables of visual variability and auditory variability, yielding three participants in each condition. Each participant was screened via verbal productions and parent report at the first experimental session in order to confirm that they did not have knowledge of 12 of the 14 possible words to be learned in the study. An additional three children were enrolled in the study but did not complete all sessions and thus were excluded from analyses, an attrition rate that is consistent with earlier work in this population (e.g., Theodore et al., 2011).

Stimuli

The stimulus set was modeled after the fourteen nouns used in Perry et al. (2010) and consisted of 14 nouns. Twelve nouns were targeted for word learning during the study period, with the other two nouns serving as alternates in the case that a child already knew one of the intended targets. The 14 nouns were real English words that were selected because it was unlikely that children in this age range would have previously acquired them. In addition, the

nouns were selected to be easily pictureable with respect to the visual exemplars. The 14 nouns used in the study are shown in the Appendix. For each of the 14 nouns, a set of auditory and visual exemplars were acquired, as described below.

The auditory exemplars consisted of verbal productions of each noun by ten different talkers (five male, five female). The talkers were monolingual English-speaking adults with no significant history of speech and/or language disorders according to self-report. Recordings were made in a sound attenuated booth. Speech was recorded via microphone (AKG D5S) connected to a pre-amplifier (Digidesign MBox2) and saved directly to hard disk. Each talker was recorded producing multiple repetitions of the 14 nouns. In addition, each talker was recorded producing the prompts for use during test phases (e.g., *Point to the bucket*). The Praat software (Boersma & Weenick, 2005) was used to excise each production of each word or prompt to an individual sound file. For each talker, one production of each target noun that was free of acoustic artifact (e.g., coughing) was selected for use in the experiment such that speaking rate (as measured by word duration) was equivalent across the 10 talkers and the 14 nouns. The selected productions of the target nouns were organized into two sets, one for use with the high auditory variability condition and one for use with the low auditory variability condition. The stimuli in the high auditory variability condition consisted of productions of nouns from all ten talkers, whereas productions of nouns from only one of the ten talkers were used as the low auditory variability stimuli. The prompts produced by the talker selected for use in the low auditory variability condition served as test stimuli for both the high and low auditory variability groups. All selected tokens were equated for root-mean-square amplitude.

The visual exemplars consisted of nine objects for each of the 14 nouns. Of the nine objects, three were for use in the low visual variability condition, three were for use in the high

visual variability condition, and three were for use during tests of generalization for both variability groups. For each noun, the three objects used in the low visual variability condition were minimally distinct from each other, and objects in the high visual variability condition were maximally distinct from each other. See Figure 1 for an example of low visual variability, high visual variability, and generalization sets for one of the selected nouns, *can*.

In order to quantify distinctions among the visual exemplars for each noun, each exemplar was categorized on five attributes that have been shown to be relevant to word learning in children including shape, color, size, material, and whether the object had an additional component that had a separate name (e.g., bucket handle, box top, zipper). Note that the first four attributes were constant across the 14 nouns, but that the last attribute was customized for each individual noun to some degree. Based on these categorizations, we calculated the number of different attributes among the exemplars for a given noun to indicate assignment to the low or high visual variability condition. For a particular noun, the exemplars used in the low visual variability group differed in 0 – 2 attributes, whereas the exemplars used in the high visual variability group differed in 4 – 5 attributes. The exemplars used for generalization differed in 3 – 4 attributes from the exemplars in either the low or high visual variability conditions.

Procedure

The experiment consisted of nine weekly sessions conducted in the laboratory. At the beginning of each session, vocabulary development was measured using the MacArthur Bates Short Form Vocabulary Checklist: Level II (Form A) of the MacArthur-Bates Communicative Developmental Inventory (CDI; Fenson et al., 1994). Parents were also contacted one month following the completion of the study to submit a follow-up CDI. During each session, the child

and parent were invited to engage in naturalistic play using the visual exemplars of their assigned condition. The children were placed at a table, either alone or on the lap of their parents, and were positioned to face the speaker at the opposite end of the table. Visual stimuli were presented to the children at midline by an experimenter sitting beside the table. Parents and experimenters were instructed to avoid saying the target nouns in order to control for the amount of times every child was hearing the words.

Sessions were broken into consecutive three-week modules that each targeted four of the twelve nouns. The order in which children were presented with each module was randomized for each child. During each session of every module, the child was trained and tested on 4 target nouns. During training, the children were encouraged to manipulate and examine the visual exemplars while the auditory stimuli were presented through the loud speaker in order to learn the name of each exemplar. Presentation of auditory exemplars was controlled by the experimenter, such that each child heard an auditory referent for each exemplar twenty times per training session. Depending on the child's assigned condition, they were either presented with twenty repetitions from one talker (low auditory variability) or two repetitions from ten talkers (high auditory variability).

During testing, learning of both trained exemplars and generalization to novel exemplars was assessed through learning trails and generalization trials. During learning trials, children were presented with two exemplars previously used during training that represented different nouns. Conversely, during generalization trials, children were presented with two exemplars that also represented different nouns but were not presented during training. For both tests, an auditory prompt directed the child to identify the noun that was named (e.g. *Point to the bucket*). Three learning trials and three generalization trials were completed for each noun and the

experimenter coded the child's response as correct or incorrect. Positive feedback was provided regardless of the child's response. All sessions were video recorded so that offline reliability analyses can be completed at a later time. Each session lasted approximately 30 minutes.

Results

Three sets of analyses were conducted in order to examine performance for the learning trials, generalization trials, and vocabulary development. Each will be addressed in turn.

Learning trials

For each child, performance for learning trials was calculated as mean proportion correct responses across the three learning trials for each noun. Mean performance across the 12 children is shown in Figure 2. Visual inspection of the figure shows suggests that both visual and auditory variability influenced learning performance. Specifically, mean proportion correct for the high visual variability group was numerically higher than the low visual variability group (0.64 and 0.55, respectively). Moreover, mean performance for the high auditory variability groups was higher than the low auditory variability group (0.62 and 0.57, respectively). Given the pilot nature of this study and the low number of participants included to date, formal statistical analyses should be interpreted with caution. However, we submitted mean proportion correct responses to between-subjects ANOVA with the factors of visual variability (high versus low) and auditory variability (high versus low). The results of the ANOVA showed a main effect of visual variability [$F(1,8) = 8.62$, $p = .019$], confirming that performance was statistically higher for those who were exposed to high compared to low visual variability. The effect of auditory variability did not reach statistical significance [$F(1,8) = 2.28$, $p = .169$], indicating that with the current sample the mean difference between the low and high auditory variability conditions is not reliable. There was no interaction between visual variability and auditory variability [$F(1,8) = 0.03$, $p = .861$], indicating that the improved performance in the high compared to the low visual variability conditions was equivalent for the low and high auditory variability conditions.

One additional analysis was performed on the learning data. Recall that on each learning trial, children were directed to indicate which of two visual exemplars matched the auditory prompt. Accordingly, chance performance during learning would result in being correct on half of the trials. In order to examine whether performance in each of the four conditions formed by crossing the two independent variables was different from chance, four one-sample t-tests were performed. The results showed that performance did not differ from chance for children in the low visual, low auditory condition [$t(2) = 0.83$, $p = .495$], but that performance in all other conditions was marginally above chance performance [$t(2) > 3.77$, $p = .064$ in all cases].

Generalization trials

Similar to that of learning trials, performance for generalization trials was calculated for each child as mean proportion correct responses across the three generalization trials for each noun. Mean performance across the 12 children is shown in Figure 3. While performance during learning trials was affected by both visual and auditory variability, the same was not found for performance on generalization trials. Specifically, mean proportion correct for the low visual variability group was numerically higher than the high visual variability group (0.63 and 0.56, respectively), which is the opposite pattern of what was observed during learning trials. However, mean performance for the high auditory variability groups was higher than the low auditory variability group (0.62 and 0.57, respectively). In other words, while auditory variability may have influenced the ability to generalize words to novel referents, visual variability did not. Again, it is important to consider the limited subject number when interpreting these findings. We submitted mean proportion correct responses to between-subjects ANOVA with the factors of visual variability (high versus low) and auditory variability (high versus low), as we did for

the learning trials. The results of the ANOVA showed that there was no main effect of visual variability [$F(1,8) = 2.26$, $p = .171$], indicating that there was no statistically significant impact on performance for those who were exposed to high compared to low visual variability. The effect of auditory variability did not reach statistical significance [$F(1,8) = .99$, $p = .349$], indicating that with the current sample the mean difference between the low and high auditory variability conditions is not reliable. Additionally, there was no interaction between visual variability and auditory variability [$F(1,8) = 0.93$, $p = .363$], indicating that the improved performance in the high compared to the low auditory variability conditions was equivalent for the low and high visual variability conditions.

Similar to that of the learning data, an additional analysis was performed on the generalization data. Because each child was instructed to pick one of two exemplars corresponding to the noun in the auditory prompt, chance performance during generalization, like in learning, would be 0.50 proportion correct. To determine whether performance in each of the four conditions formed by crossing the two independent variables was different from chance, four one-sample t-tests were performed. The results showed that performance did not differ from chance for children in both the high visual, low auditory [$t(2) = 0.98$, $p = .429$] and high visual, high auditory [$t(2) = 1.25$, $p = .337$] conditions. Performance in the other two conditions, low visual, high auditory [$t(2) = 7.62$, $p = .017$] and low visual, low auditory, [$t(2) = 9.27$, $p = .011$] was above chance performance.

Vocabulary development

Recall that vocabulary performance was calculated as noun vocabulary size based on parents report using the MacArthur CDI. As described in the Introduction, Perry et al. (2010)

reported that children in the high visual variability condition showed improved gains in vocabulary size compared to those in the low visual variability condition that could not be explained by exposure to the words used in their study. Due to the limited number of participants in the current study (n = 3 in each condition), and the degree to which vocabulary size at session one differed among the four conditions, we were not able to conduct a parallel analysis. However, two of the four conditions, the high visual, low auditory and low visual, high auditory groups, did show equivalent vocabulary size at baseline which allowed us to perform a pilot analysis on this pattern over time. Note that this comparison allows us to examine whether variability in the visual or auditory domain is a better facilitator of vocabulary development because these two conditions have variability in just one of those dimensions. Figure 4 shows that for both conditions, noun vocabularies were larger at the end of the study when compared to the beginning. However, those children who had high variability in the auditory domain demonstrated greater gains in vocabulary growth than those who had high variability in the visual domain. Further, these gains were consistently demonstrated throughout the duration of the study, as well as during the time period between the last experimental session and the one-month follow-up when the children were out of the lab. These findings may suggest that variability in the auditory domain may facilitate vocabulary growth over a relatively short period of time. However, given the limited number of subjects in each condition, we were unable to compare overall vocabulary growth across all four conditions.

Summary and Conclusions

The process of pairing a visual referent with auditory stimuli is a dynamic process that has been shown to rapidly develop between the ages of 13 to 24 months of age (Nelson, 1973; Woodward, Markman, Fitzsimmons, 1994). Past work has examined the effects of variability in the auditory and visual domains on the word learning process when each domain is examined independently. More specifically, learning lexical neighbors and developing a consistent phonetic contrast across talkers has been facilitated by auditory variability, while generalization of naming to novel visual items has been facilitated by visual variability (Rost & McMurray 2009, 2010; Perry et al. 2010). However, examining the impact of variability in both domains simultaneously had not been previously explored.

The current study contributes to our understanding of the word learning process for concrete nouns that occurs during early childhood and, more specifically, how variability in both auditory and visual domains affects this process. In addition to learning words using visual exemplars that children were trained with, we also examined the ability to generalize the name of an object to a novel visual stimulus. The results of our pilot study indicate that variability in the visual domain facilitated naming of trained visual exemplars, or learning trials, but did not have an effect on generalization of this word to novel visual exemplars during generalization trials. This finding differs from results of previous work by Perry et al. (2010) where visual variability did not have an effect on learning trials but instead influenced performance on generalization trials.

Additionally, we found that there was no interaction between the auditory and visual domains for both learning and generalization. In other words, improved performance in the high compared to the low variability in one domain was equivalent for the low and high variability

conditions in the other domain. This finding suggests that the process of learning a word and generalizing that word to other instances may be domain-specific. In a paradigm where auditory stimuli are highly variable, like those in the high auditory variability condition in this study, it is possible that variability in the auditory domain does not necessarily facilitate our ability to extend our knowledge into the visual domain. In the current study, we trained some children with highly variable auditory stimuli and low variability in visual stimuli. While we provided highly variable auditory training, we were testing generalization in the visual domain so benefits of auditory variability may not have carried over to the visual domain. Generalization benefits may be domain-specific during the word learning process.

Lastly, we examined the effects of auditory and visual variability on overall vocabulary development throughout the duration of the study and during the one-month that the children were out of the lab. Due to our small sample size and varying vocabulary sizes at the start of the study within each condition, we were unable to fully examine the effects of auditory and visual variability on overall vocabulary for all four conditions. However, at baseline, vocabulary sizes were equivalent for both the high visual, low auditory and low visual, high auditory groups. This, in turn, allowed us to examine whether variability in the visual or auditory domain is a better facilitator of vocabulary development due to the variability in just one of those dimensions. In this case, auditory variability facilitated vocabulary development.

It is important to consider the clinical implications of this pilot study's findings. When learning the name for a specific visual object, it appears that learning that word is facilitated by having several exemplars that vary in visual characteristics. For typically developing children, as well as those with language deficits, this finding may influence strategies used by parents at home, as well as speech-language pathologists working to improve expressive and receptive

language capabilities. For example, instead of having similar types of commonly used objects or toys, such as buckets or spoons, caregivers and clinicians may recognize the importance of providing a wide variety of examples that all fit within those categories to broaden the child's understanding of the new word. However, more work needs to be done to identify the best way to generalize newly learned words to novel visual exemplars.

It is clear that future work is needed to gain a better understanding of how simultaneous variability in the auditory and visual domains impacts word learning and generalization in young children. Increased sample sizes can hold promising patterns of discovery and help to not only stabilize patterns across the four conditions but also control for high inherent variability of vocabulary size within each condition at the initial session. This would allow for more reliable analysis of the effect variability has on vocabulary development. Future work is also needed to examine the effects of domain-specific generalization and whether or not there is interaction between visual and auditory domains during word learning. Lastly, future work should examine the effects of auditory and visual variability on word learning in children with hearing loss. It is possible that variability in both domains could be overwhelming or, on the other hand, it could contribute to a wider understanding of concrete nouns. In the face of auditory variability, children with hearing loss may require similar visual exemplars to provide a consistent visual referent.

References

Boersma, P. & Weenink, D. (2005). Praat: Doing phonetics by computer (Version 4.3.01) [Computer program]. Retrieved from http://www.praat.org/.

Childers, J. B., & Paik, J. H. (2009). Korean-and English-speaking children use cross-situational information to learn novel predicate terms. *Journal of Child Language*, *36*, 201-224.

Clark, E. V. (1973). What's in a word: On the child's acquisition of semantics in his first language. In T.E. Moore (Ed.), *Cognitive development and the acquisition of language.* (pp. 42-86). New York: Academic Press.Delattre, P.C., Liberman, A.M., & Cooper, F.S. (1955). Acoustic loci and transitional cues for consonants. *Journal of Acoustical Society of America, 27,* 769-773.

Dromi, E. (1986). The one-word period as a stage in language development: Quantitative and qualitative accounts. In I. Levin (Ed.), *Stage and structure: Reopening the debate*. Norwood, NJ: Ablex.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M., Mervis, C. B., Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59, 1-185.

Gershkoff-Stowe, L., & Smith, L. B. (2004). Shape and the first hundred nouns. *Child Development, 75*(4), 1098-1114.

Jusczyk, P. W. (1986). Toward a model of the development of speech perception. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 1-19). Hillsdale, NJ: Erlbaum

Klatt, D. H. (1986). The problem of variability in speech recognition and in models of speech perception. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech*

processes (pp. 300-319). Hillsdale, NJ: Erlbaum

Kuhl, P. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior & Development, 6*, 263-285.

Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America, 35,* 1773-1781.

Mervis, C. B. (1987). Child-basic object categories and early lexical development. In U. Neisser (Ed.), *Concepts and Conceptual Development: Ecological and intellectual factors in categorization* (pp. 201-233). New York: Cambridge University Press.

Miller, J. L. (1981). Effects of speaking rate on segmental distinctions. In P. D. Eimas & J. L. Miller (Eds.), *Perspectives on the study of speech* (pp. 39-74). Hillsdale, NJ: Erlbaum.

Munro, M. J. (1998). The effects of noise on the intelligibility of foreign-accented speech. *Studies in Second Language Acquisition*, *20*, 139-154.

Nelson, K. E. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development, 38*, (1, Serial No. 149).

Oakes, L. M., Coppage, D. J., & Dingel, A. (1997). By land or by sea: The role of perceptual similarity in infants' categorization of animals. *Developmental Psychology*, *33*, 396-407.

Perry, L. K., Samuelson, L. K., Malloy, L. M., & Schiffer, R. N. (2010). Learn locally, think globally: Exemplar variability supports higher-order generalization and word learning. *Psychological Science*, *21*, 1894-1902.

Pols, C. W. L (1986).Variation and interaction in speech. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 140-154). Hillsdale, NJ: Erlbaum.Rost, G. C. & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science*, *12*, 339-349.

Rost, G. C., & McMurray, B. (2010). Finding the signal by adding noise: The role of noncontrastive phonetic variability in early word learning. *Infancy*, *15*(6), 608-635.

Samuelson, L. K., & Smith, L. B. (1999). Early noun vocabularies: Do ontology, category structure and syntax correspond? *Cognition, 73*, 1-33.

Singh, L. (2008). Influences of high and low variability on infant word recognition. *Cognition*, *106*, 833-870

Singh, L., White, K. S., & Morgan, J. L. (2008). Building a word-form lexicon in the face of variable input: Influences of pitch and amplitude on early spoken word recognition. *Language Learning and Development*, *4*, 157-178.

Smith, L. B. (1999). Children's noun learning: How general learning processes make specialized learning mechanisms. In Brian MacWhinney (Ed.), *The emergence of language* (pp. 277-303).

Smith, L. B. (2003). Learning to recognize objects. *Psychological Science, 14*, 244-250.

Soja, N. N., Carey, S., & Spelke, E. S. (1991). Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition, 38*, 179-211.

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*, 381-382.

Summerfield, A. Q. (1975a). Aerodynamics versus mechanics in the control of voicing onset in consonant-vowel syllables. *Speech Perception, 4*. Department of Psychology, Queen's University of Belfast.

Thom, E. E., & Sandhofer, C. M. (2009). More is more: The relationship between vocabulary size and word extension. *Journal of Experimental Child Psychology*, *104*, 466-473.

Verbrugge, R. R., & Shankweiler, D. Prosodic information for vowel identity. (1977). *Journal of*

the Acoustical Society of America, 61, [Abstract].

Woodward, A. L., Markman, E. M., & Fitzsimmons, C. M. (1994). Rapid word learning in 13- and 18-month-olds. *Developmental Psychology, 30*, 553-566.

Figure 1

Training and test exemplars used for the noun *can*. As described in the main text, children in each training group were tested on the exemplars presented during training (learning trials) and novel exemplars (generalization trials). The exemplars used in generalization trials were the same for all training groups.

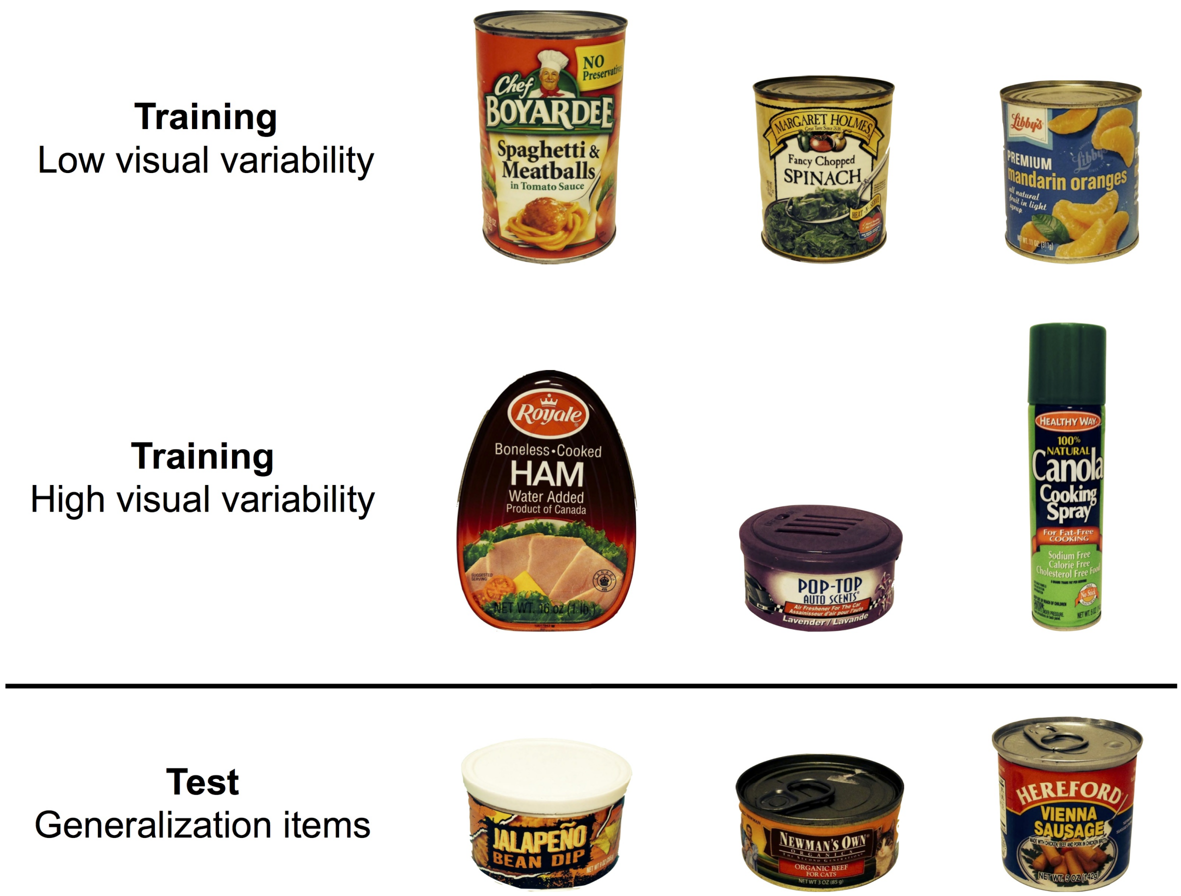


Figure 2

Mean proportion correct learning trials for the low auditory variability and high auditory variability conditions for each visual variability group. Error bars indicate standard error of the mean.

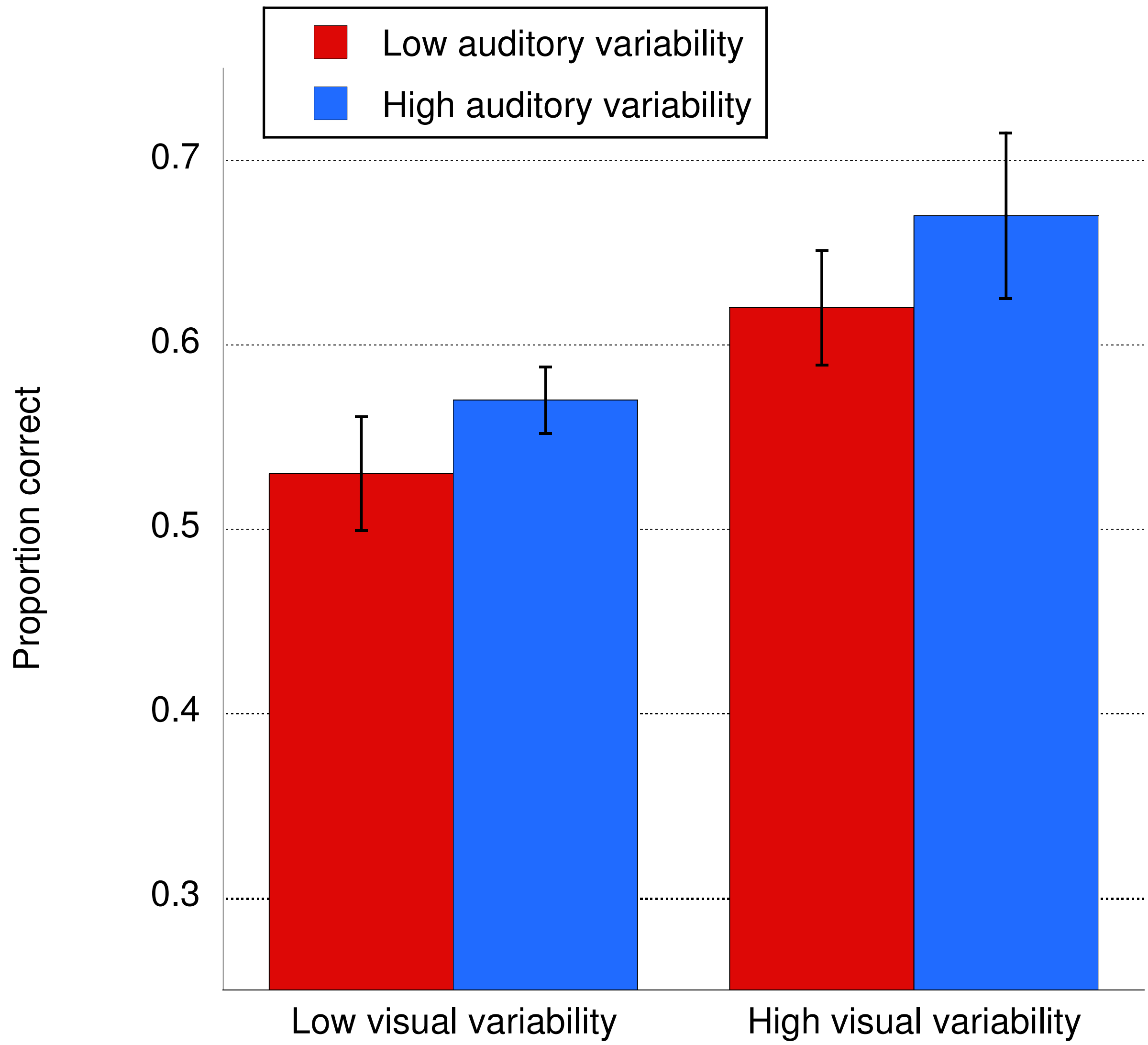


Figure 3

Mean proportion correct generalization trials for the low auditory variability and high auditory variability conditions for each visual variability group. Error bars indicate standard error of the mean.

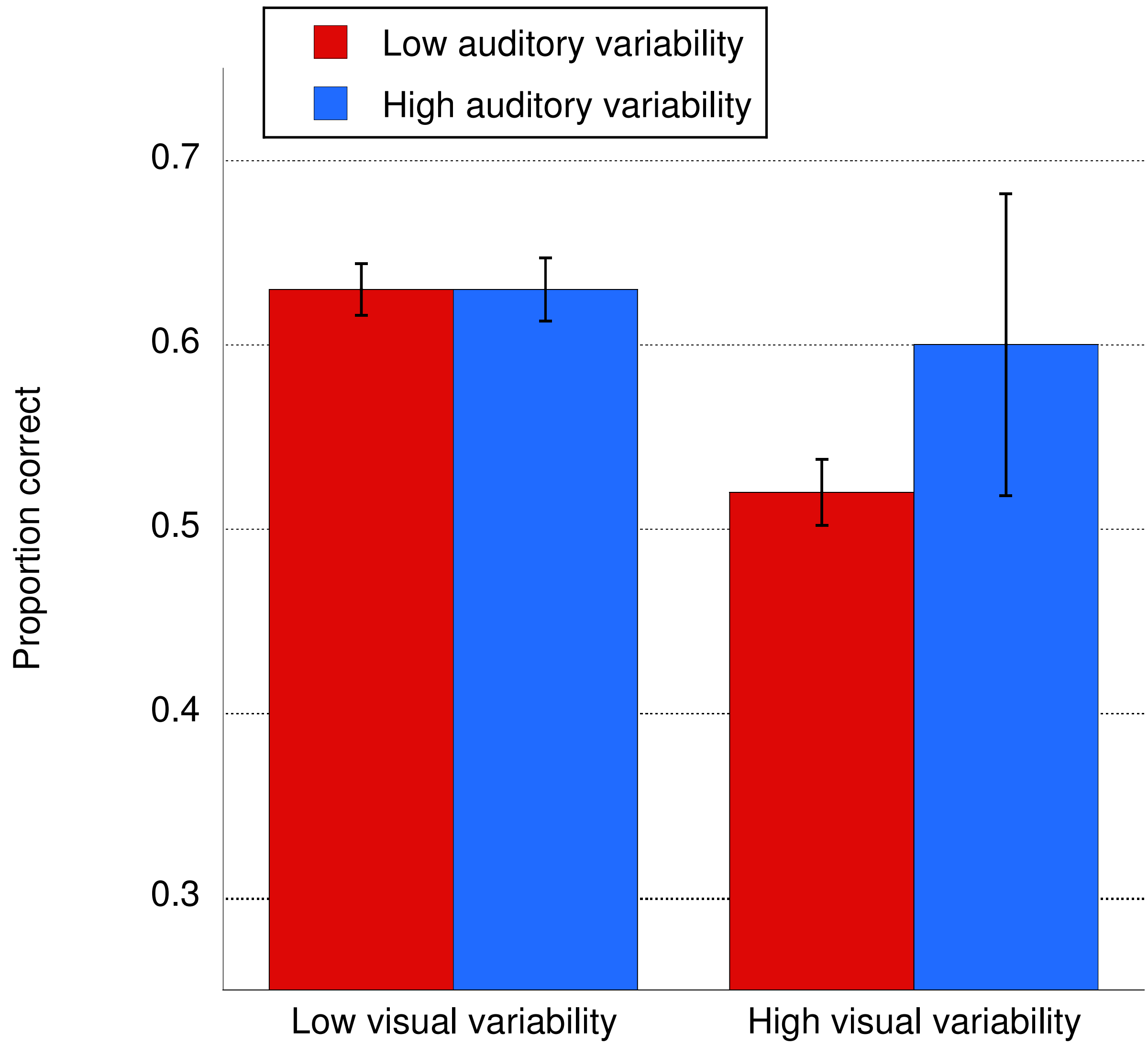
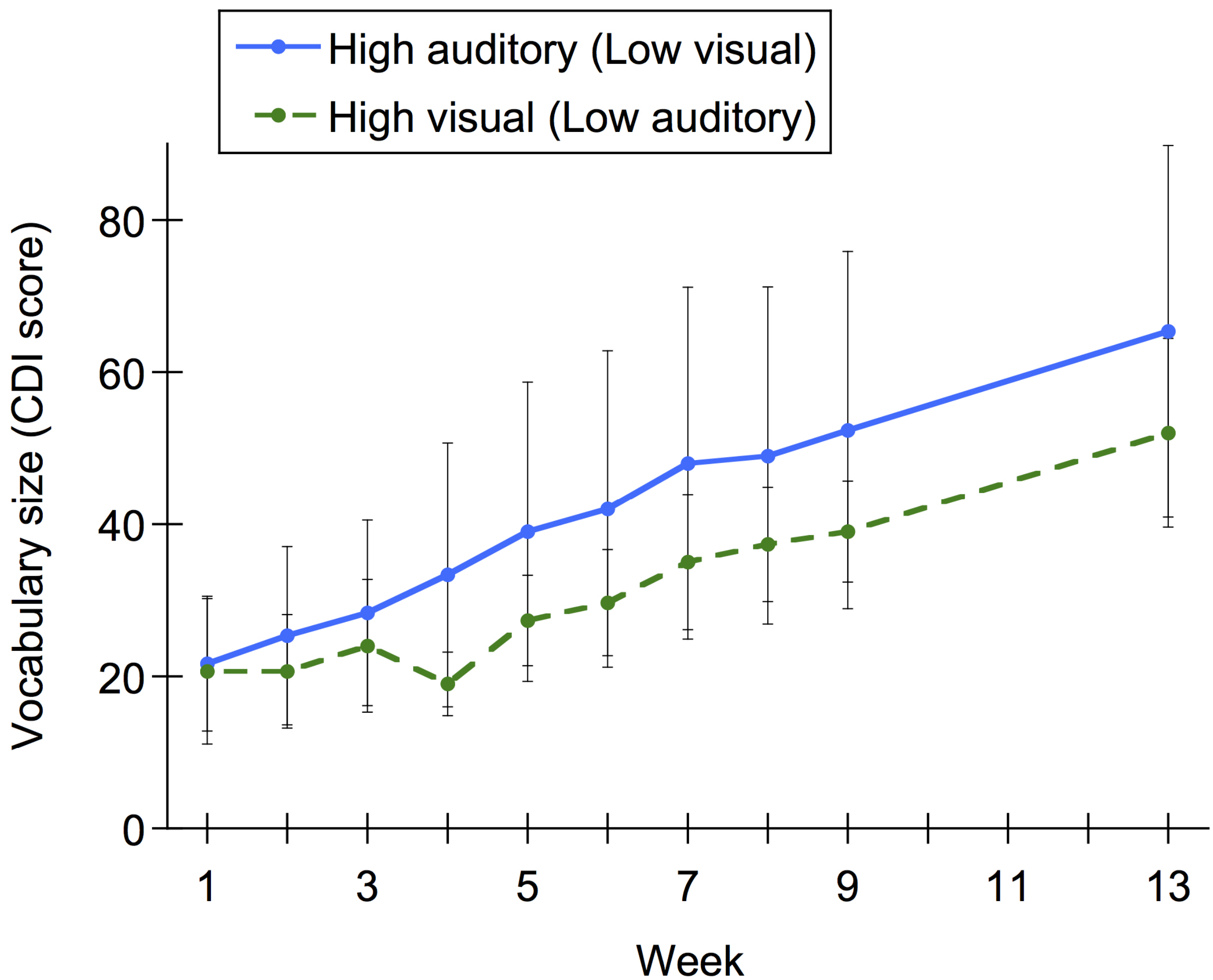


Figure 4

Mean vocabulary size for the children in the high visual, low auditory group and high auditory, low visual group spanning week 1 through the one-month follow-up.



Appendix

bead

bowl

box

bucket

can

comb

funnel

hammer

necklace

spoon

toothbrush

tractor

boot (alternate)

crown (alternate)