Infant Word Segmentation and Childhood Vocabulary Development: A Longitudinal Analysis

Leher Singh,
National University of Singapore

J. Steven Reznick, and
University of North Carolina

Liang Xuehua
National University of Singapore

Abstract

Infants begin to segment novel words from speech by 7.5 months, demonstrating an ability to track, encode and retrieve words in the context of larger units. Although it is presumed that word recognition at this stage is a prerequisite to constructing a vocabulary, the continuity between these stages of development has not yet been empirically demonstrated. The goal of the present study is to investigate whether infant word segmentation skills are indeed related to later lexical development. Two word segmentation tasks, varying in complexity, were administered in infancy and related to childhood outcome measures. Outcome measures consisted of age-normed productive vocabulary percentiles and a measure of cognitive development. Results demonstrated a strong degree of association between infant word segmentation abilities at 7 months and productive vocabulary size at 24 months. In addition, outcome groups, as defined by median vocabulary size and growth trajectories at 24 months, showed distinct word segmentation abilities as infants. These findings provide the first prospective evidence supporting the predictive validity of infant word segmentation tasks and suggest that they are indeed associated with mature word knowledge.

INTRODUCTION

In a pioneering study, Jusczyk and Aslin (1995) demonstrated that infants begin to segment words from fluent speech between 7 and 8 months of age. After familiarizing infants with a pair of novel words, such as 'cup' and 'feet', infants were then tested on recognition of those words in the context of passages. Results showed that infants listened longer to passages containing familiarized words compared with those that contained novel words. This constituted the first laboratory demonstration of emergent word knowledge in infants as young as 7 months, several months before words are associated with meaning with any regularity. This demonstrates an impressive ability in such early learners, but subsequent modifications of this paradigm have revealed limitations on infants' performance at this age. Specifically, when words change in emotion, talker gender, or fundamental frequency, infants are no longer able to reliably recognize words (Singh, Morgan & White, 2004; Houston & Jusczyk, 2000; Bortfeld & Morgan, 2010), reflecting a lack of certainty about how sound changes interact with changes in meaning. It should be noted that infants demonstrate word recognition at 6 months for words that are familiar (e.g., 'mommy') and

Address for correspondence: Leher Singh, Dept of Psychology, National University of Singapore, Block AS 4, Level 2, 9 Arts Link, Singapore 117570.

reliably generalize at 7.5 months across pitch changes for such words (Bortfeld, Rathbun, Morgan & Golinkoff, 2003; Mandel, Jusczyk & Pisoni, 1995; Singh, Nestor & Bortfeld, 2008) suggesting increased precocity for particular privileged items in the input. Other early limitations include the finding that infants are only able to segment words that conform to the predominant stress pattern of their language (Jusczyk, 1997). Later, by 11 months, infants appear to have outgrown many of these limitations and are able to recognize words independent of the surface form they assume (Bortfeld & Morgan, 2010; Houston & Jusczyk, 2000; Singh, et al., 2004). By the same time, they are able to segment words that exemplify both predominant and uncommon stress patterns (Jusczyk, Houston & Newsome, (1999). Therefore, the period between 7 to 11 months reflects a series of transitions in how word-forms (defined as phonological units that are later associated with lexical entries) are represented in memory.

These findings simultaneously reveal impressive capability and serious limitations on early word recognition at this stage. Seven to eight months of age clearly represents an important period of transition for language learning as the ability to track spoken words and to detect repetitions of those words in different contexts is likely to be a prerequisite to later mapping those words onto meaning. Without being able to perceive recurrences of words as lexically equivalent, word-meaning mapping is rendered impossible. Although it has therefore been presumed that word segmentation in infancy is directly relevant to, and necessary for, lexical development, an empirical association between infants' word segmentation abilities and later lexical development has not yet been established in a prospective design. It is not yet evident how successful word segmentation in infancy might bear on word learning later in early childhood and perhaps more critically, how limitations in early word segmentation might impact upon the subsequent emergence and expansion of a vocabulary.

As is common in infant language research, conclusions about infant word segmentation abilities were drawn from central tendencies derived from group data. However, an equally common feature of such research is that there is a high degree of variability about the mean. Such variability is typically disregarded as reflecting sampling error or other uninformative sources of variation, and individual data points are rarely analyzed independently. However, it is highly possible that this dispersion of data points about the mean, if reliably observed across trials and tasks within participants, could constitute relevant data that draws meaningful distinctions within a sample. The goal of the present study is to determine whether individual differences in infant word segmentation relate to word learning later in childhood. In addition, the current study aims to assess the construct validity of early word recognition tasks to determine whether they indeed predict the infant's later facility with mapping words to meaning.

The Role of Word-form Recognition in Language Development

A tacit assumption in the infant word recognition literature is that experimental tasks are designed to probe the precursors to word knowledge. However, the nature of the relationship between infant word recognition and later language development is not by any means obvious. By necessity, infants must be able to understand what qualifies as a word in their language prior to the assignment of meaning. This is not a trivial task for young infants as languages of the world differ in how they exploit sound to communicate meaning. The sequence of events involved in word-form recognition as it is traditionally assessed in the laboratory is therefore relatively complex. Infants must first detect repetitions of words in citation form, parse subsequent test passages and segment repetitions of words from those passages. Additionally, the scope of phonological representation is a fine balance. On one hand, memories for words have to be sufficiently broad to incorporate the rampant variability inherent in natural discourse and to normalize for accents, different voices, emotions, and other factors. On the other hand, memories have to be sufficiently specific to

not incorrectly equate minimal pairs or tolerate mispronunciations. Therefore, learners have to negotiate a tension between generality and specificity of representation in a way that allows their phonological definition of a word to correspond precisely to its semantic definition.

The exact mechanisms by which word form representations culminate in a vocabulary remain undefined. However, there are identified candidate factors that facilitate the formation of robust word form representations. These factors include frequency in the input (Bortfeld, et al., 2003; Mandel, et al., 1995), pre-exposure to word forms prior to a word learning task (Swingley, 2007), inclusion of word-like units in varying frames (Gomez, 2002), and the inclusion of the prosodic undulations associated with infant-directed speech (Ma, Golinkoff, Hirsh-Pasek & Houston, 2011). Furthermore, it is possible that word-form recognition builds knowledge beyond the lexicon, such as phonetic category structure acquisition (Swingley, 2009), detection of statistical regularity, and classification of words into syntactic categories (e.g. Mintz, 2003; Shi & Gauthier, 2005). Although these areas have not traditionally been investigated in tandem, word-form recognition may be both an outgrowth and a cause of development in other areas of language. Additional empirical evidence is required to verify this possibility. To summarize, it is hard to dispute the necessity of stable word-form representations for later word-meaning mappings. However, many open questions remain about the nature of the link between these two abilities, particularly with regards to precisely which drivers of growth in infant word knowledge may stimulate growth in later language acquisition.

Predicting Child Language Development from Infant Behavior

There have been several initiatives to link basic auditory responsiveness (both to speech and non-linguistic auditory stimuli) to later language development. In this line of inquiry, Benasich and Tallal (2002) found that discrimination of brief, rapidly presented auditory stimuli at 7.5 months predicted size of both receptive and expressive vocabulary at 12, 16, 24 and 36 months. Additionally, performance at 7.5 months on measures of rapid auditory processing (RAP) and correlated vocabulary sizes at 12-36 months successfully discriminated between infants with and without a family history of Specific Language Impairment (SLI). Trehub and Henderson (1996) found that infants' temporal resolution ability at 6.5 months (as measured by detection thresholds of auditory gaps of 12-40 milliseconds between pure-tone stimuli) predicted size of expressive vocabulary at 16-29 months.

Furthermore, neonatal speech discrimination predicts later language abilities as demonstrated in a number of reports by Molfese & Molfese (e.g. Molfese & Molfese, 1985, 1997; for a review, see Molfese, Molfese & Molfese, 2007). These studies have demonstrated that individual differences in neonatal event-related potential waveforms bilaterally in response to consonant sounds predict language development at 3 years of age (Molfese & Molfese, 2005), as well as IQ scores at age 5 (Molfese & Molfese, 1997). In addition, literacy skills were predicted by infant responsiveness to consonant contrasts with newborn ERP components successfully discriminating between poor and typical readers at age eight years. These studies investigate the associations between a basic responsivity to both linguistic or non-linguistic stimuli and later language skills. It appears that a neurophysiological receptivity to speech and to non-speech predicts later oral and written language skills in multiple domains, raising valuable questions about any observed associations in the present study.

More recently however, there have been a series of studies to investigate infant speech processing at a point where infants are known to have developed language-specific knowledge (Jusczyk, 1997; Kuhl, 2004), capitalizing on infants' acquisition of aspects of

native language structure to generate predictor variables. At the level of the phoneme, Tsao, Liu and Kuhl (2004) demonstrated that phoneme discrimination of a non-native vowel contrast (that corresponded to different phonetic categories in English) at 6 months correlated with productive vocabulary size at 13, 16 and 24 months. In a follow-up to this, Kuhl, Conboy, Padden, Nelson & Pruitt (2005) investigated the relationship between non-native phoneme discrimination at 7 months - a Mandarin fricative-affricate distinction - and later language outcomes. In addition, a single native contrast was included as a predictor variable. Infants' abilities to discriminate the native and non-native contrasts employed in this study were correlated with vocabulary sizes at 18 and 24 months, as well as with reduced sentential complexity at 24 months. In addition, accuracy of discrimination of native contrasts at 7 months correlated negatively with accuracy of discrimination of a non-native contrast at the same age, suggesting that the transition to language-specific attunement is associated with reduced sensitivity to the non-native contrast employed in the study. An improvement in discriminating the native contrast used as well as a concurrent decline in discriminating the non-native contrast were both associated with positive language outcomes at 24 months.

An interesting finding from Tsao, et al. (2004) was that although the primary dependent measure in phoneme discrimination tasks is accuracy of discrimination (percent correct), the number of trials to reach criterion emerged as a more reliable predictor of later language than accuracy of discrimination. This suggests that a primary source of continuity in this study may have been the speed with which the phonetic contrast presented during the task was processed, rather than the stock of linguistic knowledge participants had accumulated.

Longitudinal investigations have since been initiated to investigate larger-scale units of the language code. With respect to word recognition (lexical access), Fernald, Perfors and Marchman (2006) found that individual differences in lexical access at 25 months are predicted by productive vocabulary size, as estimated by the MacArthur-Bates Communicative Development Inventories, between 15 and 25 months. Again, as in the case of phoneme discrimination, the speed with which learners launched a correct eye-movement towards a labeled target proved to be more strongly associated with vocabulary size at 25 months than accuracy of comprehension. In a follow-up to this study, Marchman and Fernald (2008) demonstrated that the efficiency of on-line processing of their stimuli at 25 months was a strong predictor of cognitive and linguistic outcomes at 8 years of age. Together, these studies suggest that at phonetic and lexical levels, the efficiency with which incoming information is processed, as well as the accuracy of processing, predict formal language skills.

In addition to phoneme discrimination and lexical access, further studies have investigated the relationship between neurophysiological correlates of prosodic stress and later language outcomes. In a cohort of 5-month-old German speaking infants, Weber, Hahne, Friedrich and Friederici (2005) investigated mismatch negativity responses to dominant (trochaic) versus non-dominant (iambic) stress patterns in the native language using an oddball paradigm in which a pair of contrastive stimuli are presented in the same test session and responses to each stimulus type are measured. Dependent measures from this task were associated with productive vocabulary at 12 and 24 months. Findings demonstrated that infants with low production outcomes at both age groups showed a reduced amplitude of response in the stress-pattern discrimination task, suggesting that reduced sensitivity to predominant stress patterns in the input is associated with risk for later language impairment. In a related study, Cristia and Seidl (2011) demonstrated that prosodic sensitivity at 6 months predicted vocabulary size as well as use of inflectional morphology at 18 months, demonstrating interactive relationships between tiers of the language code in a longitudinal design.

The most significant implication of these research programs is the identification of relevant aspects of infant speech processing that predict later performance in the domains of phoneme discrimination, lexical access and prosodic stress sensitivity. However, the success of these studies also invites further investigation at the level of the word, where there is an equally rich infant literature that suggests expected developmental milestones. Closest to the current investigation is a study investigating word segmentation, as well as other types of infant tasks, as they relate to later formal language skills. In a retrospective design investigating several aspects of infant language processing (e.g., language discrimination, word segmentation, use of phonotactic cues in segmentation) as they relate to later language skills, Newman, Ratner, Jusczyk, Jusczyk and Dow (2006) sought to connect archival data from infant studies to language abilities in childhood. Among the predictor variables for which archival data were obtainable, the authors found the strongest relationships between infancy and childhood to emerge for word segmentation measured between 7.5 and 12 months and an overall language quotient (including semantic and syntactic subscales) at 4 to 6 years. In their analysis, Newman and colleagues examined the subset of children who fell into the tails of the distribution at outcome (i.e., those learners below the 15th percentile or above the 85th percentile), finding significantly different levels of success in infant word segmentation tasks based on whether learners were categorized at the highest or lowest extreme of the distribution at 24 months. Interestingly, Newman and colleagues found no significant differences in outcome IQ (either verbal or non-verbal components) for children who segmented words successfully compared to those that did not, pointing to domain-specific continuity between predictor and outcomes variables. Examining the extremes of the distribution is an approach that establishes very gross distinctions at outcome by including those infants who fall into either extreme, thus disregarding 70% of usable data points. This excludes the majority of participants, but it may hold great potential in identifying clinical sub-groups within the sample. Statistically, evidence of group differences between the upper and lower extremes is likely to yield more pronounced effects than an analysis that incorporates the entire distribution of outcome values. In their Discussion, Newman and colleagues acknowledge the limitations of their retrospective design, such as selection bias at outcome and the availability of binary prediction data, stating that "The current study suffers from a number of limitations that suggest the strong need for prospective study of the potential relationships between early speech segmentation abilities and later language development." The present study is designed to address this need and to probe relationships between word segmentation and vocabulary development in a prospective design, as previously evinced for phoneme perception, prosodic stress discrimination and lexical access.

Overview of Experimental Design—In the current study, word segmentation skills were assessed in a cohort of infants at 7.5 months. In an initial task, word segmentation at 7.5 months was assessed using a paradigm highly similar to that of Jusczyk & Aslin (1995), where infants were familiarized with two words and then tested on recognition of those words in the context of passages. As in Jusczyk & Aslin's study, an increase in listening time to passages containing familiarized words relative to novel passages constituted evidence of word recognition. In a second session, a more complex task was administered. This task is equivalent to the task employed by Singh, et al. (2008a): Infants were familiarized with two words, one of which was in a matched pitch to the target word in recognition passages and one of which was in a mismatched pitch to the target word in the recognition passages. Therefore, this experiment generated two values: Recognition of the matched word and a more advanced ability to generalize and recognize the pitch-mismatched word. Infants' vocabulary size was then monitored from 8 to 24 months on a bimonthly basis. The Bayley Scales of Infant Development (BSID, 2nd Ed) were also administered to derive a Mental Development Index at 23 months. The overarching goal of

the study was to explore potential relationships between infant word segmentation, productive vocabulary and cognitive skills at outcome (23-24 months).

Methodology

Participants

Forty typically developing infants (22 girls and 18 boys) participated in the study and were tracked from 7.5 months through 24 months. There were no known medical conditions within the sample. All infants were full-term with the exception of two, who were born at 37 and 35 weeks. The average age of infants at intake was 221 days (range: 213 to 242) and at outcome was 24 months and 8 days (range: 24 months and 1 day to 24 months and 13 days). Participants were recruited from the Boston area and all measures were collected in the laboratory. Four infants were raised in bilingual homes, with primary emphasis on English, but all of the others were raised in solely English-speaking homes.

Procedure

Infant Word Segmentation Tasks—Infants were tested using the Headturn Preference Procedure (HPP) (Kemler-Nelson, Jusczyk, Mandel, Myers, Turk, & Gerken, 1995), which was implemented identically to previous studies (Bortfeld et al., 2005; Jusczyk & Aslin, 1995; Singh et al., 2004; Singh et al., 2008a, Singh, et al., 2008b). The infant was seated on the parent's lap facing a center light. Each trial began with the center light flashing until the infant fixated on the flashing light. Then this light was turned off, and one of two side lights began to flash to attract the infant's attention. When the infant fixated the side light, speech stimuli for that trial began to play. The sound and light remained on for the duration of the infant's fixation.

In one session, familiarization began with trials alternating between the two target words in citation form (either 'bike' and 'hat' or 'tree' and 'pear') until 30 seconds of familiarization had accrued for each word. Henceforth, these two target words will be referred to as 'familiarized words'. Half of the sentences contained the two familiarized words and half contained the two non-familiarized words. Recognition testing consisted of four blocks of trials, each block containing one trial with each of the four passages. Each passage comprised sentences containing one of the four words listed above, two of which were familiarized words. Thus, in the recognition phase, infants heard four passages, two of which comprised sentences containing the familiarized words. The order of passages within each block was randomized for each infant, as was the order of sentences within passages on each trial. Word recognition was indexed by the amount of time infants listened to passages containing familiarized words versus non-familiarized words during the recognition phase.

All words and passages are listed in Appendix A. Stimuli were identical to those used in previous studies (Singh, Morgan & White, 2004; Singh, et al., 2008a; Singh, 2007; Singh, et al., 2008b) and are therefore known to generate reliable results in the current paradigm at 7.5 months. Listening times to passages containing novel words were subtracted from those containing familiarized words to generate a recognition score for both words. When an increase in listening time to familiarized words compared to listening times to non-familiarized words was significantly different, this is taken as evidence of successful word recognition.

In a variation of this task, infants were tested on a similar paradigm in which they were expected to generalize across two distinct exemplars of a word. Previous research employing one familiarization target that matches in indexical detail between training and test and one word which mismatches between training and test has demonstrated that infants encounter

difficulty with the mismatched form (Houston & Jusczyk, 2000; Bortfeld & Morgan, 2010, Singh, Morgan & White, 2004; Singh, White and Morgan, 2008). This task was included because it probes a more sophisticated form of word recognition – generalization across distinct surface forms – using an experimental design that has been corroborated across different laboratories and different sources of indexical detail. In the current study, pitch variation was incorporated into the task to produce a mismatch between one target word between training and test. In a design identical to that of Singh et al. (2008a), one word was presented in a relatively high pitch and the other in a relatively low pitch. Passages containing the familiarized words as well as novel passages were presented in a high pitch for half of the infants and in a low pitch for the other half of the sample. One set of stimuli (henceforth the High Pitch set) was created by raising the fundamental frequency of all words and passages by ¼ octave (3 semitones). This was done by applying a uniform translation of all pitch points up by ¼ octave. A second set of stimuli (henceforth labeled the Low Pitch set) was created by decreasing the fundamental frequency of all words by the same amount (¼ octave). Therefore, the difference between the two sets of stimuli was half an octave. Both sets of stimuli involved pitch manipulations, so that infants' preferences would not be affected by the naturalness of the stimuli. All acoustic transformations and measurements were done using a script in the Praat program (Boersma & Weenink, 1996). Amplitude and duration measures were identical across the two stimulus sets. Stimuli are identical to those used in Singh, White and Morgan (2008) where acoustic analyses are further detailed. Each participant heard a different pair of words across simple and complex segmentation tasks.

During familiarization, infants heard two words produced in citation form. Half of the infants heard the words *bike* and *hat,* and the other half heard *tree* and *pear.* For each infant, one word was heard in High Pitch and the other in Low Pitch. For half the infants, the High Pitch word was the matched familiarization item and the Low Pitch item was the mismatched familiarization item. For the other half of the infants, the Low Pitch word was the matched familiarization item and the High Pitch word was the mismatched familiarization item. During recognition testing, infants heard passages containing all four words. Half the infants entering the study heard all passages in a high pitch and half heard all passages in a low pitch. In previous studies, Singh, et al (2008b) found statistically reliable effects for recognition of pitch-matched words but not for pitch-mismatched words in a different sample of infants at the same age (7.5 months).

To maximize the representativeness of infants' data, parents were asked to bring their infants to the laboratory during the infant's most alert hours, and if infants appeared inattentive during the task, parents were asked to reschedule for another time to maximize the possibility of 'clean' data that best represented their infants' abilities. Eighty-seven percent of infants were tested on each task on different days and the other 13% percent were tested on the same day. The order of testing was counterbalanced across infants and the average interval between testing sessions was 4 days (range: 1 to 7 days).

Both segmentation tasks require a similar set of cognitive and linguistic abilities, although there are some important distinctions. Successful segmentation in the simple task entails encoding an isolated word-form, segmenting sentences in test passages, and comparing the product of segmentation with the familiarized information. In the complex task, this comparison becomes more challenging as it involves matching segmental detail but failing to match suprasegmental details. It remains debatable whether this process engages top-down or bottom-up processes (or both) (Aslin, Woodward, LaMendola & Bever, 1996; Brent, 1999), but it is possible that these tasks are distinguished primarily based on processing burden. Potential evidence for this hypothesis could be generated by investigating whether generalization across varying forms is easier for infants when the

segmentation burden is lightened (e.g., in phrase-final position). Alternatively, it is quite possible that simple and complex segmentation tasks engage a distinct set of cognitive processes. For example, simple segmentation tasks may involve more targeted attention to similarity across stimuli. By contrast, complex segmentation may involve attentiveness to similarity as well as inhibiting attention to surface differences (e.g., pitch). This hypothesis can be empirically tested by investigating the relationship between collateral tasks involving focused attention and inhibition, and both types of segmentation tasks.

Vocabulary Measures—Between 8 and 24 months, the MacArthur-Bates Communicative Development Inventories (MCDI) were administered on a bimonthly basis to provide a parental estimation of vocabulary size (Fenson, Dale, Reznick, Thal, Bates, Hartung, Pethick & Reilly, 1993). The Infant version of the MCDI was administered from 8 to 14 months and the Toddler Version from 16 to 24 months. The primary dependent measures of interest in the present study were productive vocabulary size at each point of assessment and productive vocabulary percentile at 24 months (Fenson, et al., 1993).

Cognitive Measures—At 23 months, the Mental Development Index of the Bayley Scales of Infant Development (2nd Ed.) (BSID-II) (Bayley, 1993) was administered. The goal of the BSID-II was to arrive at a measure of cognitive abilities as represented by the Mental Development Index (MDI). The MDI in infancy has been shown to be useful for assessing infant cognitive development due to its psychometric properties and extensive standardization (Colombo, 1993). However, the Bayley MDI also correlates with concurrent language abilities (e.g. Rescorla & Alley, 2001) suggesting that it is not a language independent measure of cognitive functioning. The Bayley MDI has been widely used as an outcome measure in children in longitudinal studies of language development in typically and atypically developing populations (Rescorla, 1998; 2009; Rescorla, Dahlsgaard & Roberts, 2000; Zwaigenbaum, Bryson, Lord, Rogers, Carter & Carver, 2009). While the Bayley MDI is regarded to be a good outcome and concurrent measure of cognitive functioning in toddlers, it should be noted that its predictive validity with later intellectual functioning is questionable (Gibbs & Teti, 1990; Whatley, 1987).

Results

Infant Word Segmentation

Simple Segmentation.

In this task, infants were familiarized with two words and then tested on recognition of those words in the context of passages. As in previous investigations using this procedure (Singh, 2008; Singh, et al, 2004; 2008a; 2008b), recognition scores were calculated by subtracting listening times to passages containing familiarized words from those containing non-familiarized words[1]. As there was no effect of item and no manipulation of familiarization stimuli, recognition scores for both words were averaged, as in Jusczyk & Aslin's (1995) study.

---

[1]The justification for using recognition scores in this and prior investigations is based on interpretability and data reduction. Infant data are typically high in variance. One source of variance in data is that some infants tend to listen more (or less) to *all* the stimuli. By subtracting listening times to passages with non-familiar words from listening times to passages with familiar words, it is possible to remove a portion of that variance. Statistically, comparison of listening times to familiarized passages to non-familiarized passages is equivalent to comparing recognition scores to zero. Therefore, in terms of interpretability, it allows for reduction of two factors to a single factor without altering the statistical outcome in any way. Another option would have been to use preference quotient as employed in previous studies (e.g. Houston, Horn, Qi, Ting & Gao, 2007). However, our overall pattern of results did not change with this dependent variable and the quotients correlated highly with our recognition scores for all tasks. We have therefore retained recognition scores in order to maintain consistency with previous studies on which this study was based.

Recognition scores and standard deviations for simple segmentation tasks are displayed in Table 1. Recognition scores differed significantly from zero, $t(39) = 4.98$, $p<.001$ (two-tailed). This demonstrates that on average, infants recognized familiarized words in the context of passages. These findings are consistent with those of Juszcyk & Aslin (1995).

Infant Word Segmentation

Complex Segmentation

In this task, infants heard one word in a pitch-matched context and one word in a pitch-mismatched context. Recognition scores were collapsed over match type (i.e., low pitch word/low pitch passages vs. high pitch word/high pitch passages) as there was no effect of passage type on recognition scores, nor was there any effect of item herein or in the previous instantiation of this method (Singh, et al., 2008a). Recognition scores for pitch-matched and -mismatched words were averaged to generate a composite for the complex segmentation task. In addition, recognition scores were calculated separately for pitch-matched and -mismatched words as the ability to recognize a word while generalizing to an acoustically dissimilar exemplar (pitch mismatched words) is a more demanding task than recognizing a word matched in pitch. All values are reported in Table 1. The capacity for recognition and generalization has been reported to emerge at 9 months using these stimuli in the same laboratory setting (Singh, et al., 2008a). Individual comparisons of recognition scores in the current sample revealed that on average (combining both pitch-matched and -mismatched items), recognition scores departed significantly from zero, $t(39) = 4.83$, $p<.001$. When comparisons were computed for pitch-matched and -mismatched items separately, recognition scores differed from zero for pitch-matched items $t(39) = 5.39$, $p<.0001$ but not for pitch-mismatched items, $t(39) = 1.04$, NS, consistent with prior cross-sectional manipulations of indexical properties, such as pitch, emotion and talker gender. Findings for both simple and complex tasks provide a replication of previous findings in infants from 7-8 months (Juszcyk & Aslin, 1995; Singh, et al., 2004; 2008).

Outcome Measures

Vocabulary measures were tracked from 8 to 24 month using the MCDI on a bimonthly basis. Mean productive vocabulary size for each participant is plotted in Figure 1a and means and standard deviations at each stage are reported in Table 2. In addition, the Bayley Scales were administered at 23 months and are reported in Table 2. The MDI was significantly correlated with outcome vocabulary size percentile, $(r(38) = .67$, $p<.0001)$, consistent with previous demonstrations of the Bayley MDI correlating with MCDI measures (Dale, Bates, Reznick & Morrissett, 1989). In addition, there was an exact overlap between infants who were above and below the 50th percentile for the MCDI as those who were above and below the 50th percentile on the Bayley Scales with identical group membership for each instrument.

Relationships between Infant Word Segmentation Tasks and Outcome Measures

An initial set of analyses explored correlations between infant measures and outcome vocabulary size (see Table 3). Recognition scores for the simple segmentation task correlated with productive vocabulary size at 24 months ($r(40) = .32$ $p<.05$) (see Figure 2a for scatterplot). Mean recognition scores for complex segmentation tasks were also significantly correlated with productive vocabulary size at 24 months, ($r(40) = .51$, $p<.01$ (see Figure 2b for scatterplot). Within the complex task, correlations were significant for matched items ($r(40) = .4$ $p<.05$) and only marginally significant for mismatched items ($r(40) = .29$ $p=.07$). Finally, simple segmentation scores were significantly correlated with the Bayley MDI at 23 months MDI, $r(40) = .35$. $p<.05$ (see Figure 2c) but complex segmentation scores were not. These initial analyses confirm a relationship between both

types of infant word segmentation tasks and outcome vocabulary size, and between simple segmentation scores and Bayley MDI scores.

Recognition scores were the dependent variables of greatest interest, but attentional measures (i.e., total listening times and number of trials required to complete familiarization) were included as secondary measures. The two attentional measures obtainable from our task were mean listening time during the test portion of the experiments and the mean number of trials required to attain familiarization during the training portion of the experiments. Group means are reported in Table 1. Although listening times during test did not correlate with outcome vocabulary size, there was a significant negative correlation between the number of familiarization trials required and outcome vocabulary size, r(38) = −.37, p<.05 (see Table 3).

Latent Growth Curve Modeling of Outcome Groups

Further analyses were aimed at examining the relationship between infant segmentation measures and growth trajectories in vocabulary size between 8 to 24 months using latent growth curve modeling. Growth trajectories for the entire sample are plotted in Figure 1a. Figure 1a shows considerable within-sample variation in growth trajectories: some participants demonstrated more gradual growth, which may be better captured by linear or quadratic models while for those with more rapid vocabulary growth patterns, higher order models may provide a better fit. Figures 1b and 1c depict vocabulary growth over time for children who were below and above the median at outcome (24 months) respectively. High vocabulary outcome infants refer to those with productive vocabulary sizes above the 50th percentile at 24 months and low vocabulary outcome infants refer to those below the median at 24 months. High and low vocabulary children did not differ in vocabulary size initially (at 8 months), but the two groups began to differ significantly in productive vocabulary size at 12 months, $t(38) = -2.87$, $p<.01$ and continued to differ through 24 months. These groups are necessarily distinct at 24 months, but it is not clear whether they are distinct in vocabulary growth patterns over time. To investigate this, the growth data were divided into high and low vocabulary outcome groups (above or below the 50th percentile at outcome) and growth models were fit to the two datasets to determine whether the two groups are distinct in terms of underlying growth patterns. As a first step, a growth model was determined for the entire sample and in a second step, the two vocabulary outcome groups were then fit to this model. Table 4 shows the results of the model selection. Four models were compared: linear, quadratic, cubic and quartic models. In Table 4, "GFI" denotes the "Goodness of Fit Index" and "Pr > Chi-Square" refers to the value, where a p value of less than .05 suggests that the model is a poor fit for the data. In the first row of Table 4, it is evident that for the entire sample, none of models are a good fit ($p<.05$ for all models). However, as suggested by Figure 1a, this may be due to the large variability in individual growth trajectories. Models were then fit to the high and low vocabulary outcome groups separately. In the second row of Table 4, it appears that linear and quadratic models are a good fit for the low vocabulary outcome group who show more gradual growth. The results from the cubic and quartic models were not included as they provide a very poor fit to the data ($p < .0001$). In the third row of Table 4, it appears that quadratic, cubic and quartic models fit the high vocabulary outcome data well, confirming observations that higher order models may better capture the more rapid growth trajectories plotted in Figure 1c while lower order models may better suit the more gradual trajectories plotted in Figure 1b.

These data suggest that infants who fall above or below the median at outcome are also distinct in terms of growth patterns. Given that these two groups are distinct in process (growth over time) and product (outcome), further analyses are aimed at determining whether differences in infant segmentation tasks are associated with outcome groups and by extension, with different growth patterns. In a 2 × 2 Group (high and low vocabulary

outcome) by Task (simple vs. complex segmentation), there was no main effect of task, no interaction of task and group but a main effect of group, $F(1,38) = 10.83$, $p < .01$, with high vocabulary outcome infants showing higher recognition scores on average compared with low vocabulary outcome groups. Separating analyses by task, for the simple segmentation task, there was a marginal effect of group, $F(1,38) = 3.52$, $p = .07$ with marginally higher recognition scores for high vocabulary outcome infants. For the complex segmentation task, there was a main effect of group, $F(1,38) = 12.56$, $p < .01$ with higher recognition scores observed in high vocabulary outcome infants compared with low vocabulary outcome infants (see Figure 3a). While causal relationships cannot be presumed, the fact that there are differences in segmentation prior to group differences in productive vocabulary size is suggestive of segmentation abilities being a potential precursor to building a productive vocabulary. In summary, although differences between complex and simple segmentation tasks were not sufficient to yield a group by task interaction, group distinctions in infancy seem to be most robust in the complex segmentation tasks.

Further analyses were aimed at comparing outcome groups on basic segmentation versus segmentation and generalization. Within the complex segmentation task, group differences were examined for matched words and mismatched words separately. There was no group (high/low vocabulary outcome) by type (matched/mismatched) interaction, $F(1,38) = .1$, NS although there was a main effect of task type, $F(1,38) = 8.35$, $p < .01$ and a main effect of group, $F(1,38) = 12.56$, $p < 01$. For matched words, there was a main effect of group, $F(1,38) = 5.02$, $p < .05$ as well as for mismatched words, $F(1,38) = 4.43$, $p < .05$ (see Figure 3b).

The previous analyses were focused on differences between vocabulary growth/outcome groups in simple and complex segmentation tasks. However, it is equally important to know whether recognition scores for each group are significantly different from baseline to determine whether both groups succeed at the tasks or whether successful word recognition is only evidenced in one group. Recognition scores for each group were therefore compared to baseline (zero) to determine whether both groups showed equal success in word recognition across tasks. Recognition scores for the simple segmentation task were significantly different from zero for both the low and high vocabulary groups, $t(18) = 2.33$, $p < .05$ and $t(20) = 4.76$, $p < .05$ respectively, showing successful word recognition on the part of both groups in this task. However, average recognition scores for the complex task were significantly different from zero for the high vocabulary group, $t(20) = 7.45$, $p < .001$ but were not significant for the low vocabulary group, $t(18) = 1.08$, NS. To determine whether these group differences were attributable to the difficulties in generalizing across the mismatched words, the significance of recognition scores was assessed for matched and mismatched words separately for high and low vocabulary groups. Results showed that for matched words, both low and high vocabulary groups demonstrated significant recognition scores, $t(18) = 2.09$, $p = .05$ and $t(20) = 6.1$, $p < .001$. However, for mismatched words, the high vocabulary group showed significant recognition scores, $t(20) = 2.053$, $p = .05$ but the low vocabulary group did not, $t(18) = -.87$, NS. This suggests that children in the high vocabulary outcome group were able to generalize across acoustically dissimilar forms as infants, distinguishing themselves from their low vocabulary counterparts as early as 7-8 months. This finding also suggests that previous evidence of 7.5-month-old infants inability to generalize across word forms may not capture the developmental profile of all infants at this age, but rather primarily those who attain a lower vocabulary size as toddlers.

In addition to sorting the data based on MCDI outcomes, it is possible to sort the data based on infant measures and establish whether group differences exist in outcome vocabulary size. We therefore sorted the data based on whether infants were above or below the median on the simple and complex segmentation tasks and compared vocabulary sizes across groups for each infant task. When infants are grouped based on whether they fell above or below

the median recognition score on the simple segmentation task, there was no difference in vocabulary size at 24 months, $t(19) = -1.49$, NS. By contrast, when infants were grouped based on whether they fell above or below the median on the complex segmentation task, infants who were above the median had significantly larger vocabulary sizes at 24 months in comparison to those who were below the median, $t(19) = -3.9$, $p<.001$.

Discussion

The goal of this study was to investigate relationships between infant word segmentation tasks to vocabulary development in childhood in a prospective design. The present results provide strong evidence for this association, and support the kind of continuity from infancy through early childhood previously reported at the level of the phoneme (Tsao, et al., 2004) and in speed of lexical access (e.g. Fernald, et al., 2006) in studies employing a comparable prospective, longitudinal approach. The current study complements the findings of Newman and her colleagues (2006) in attesting to infant word segmentation tasks as potential predictors of later language outcomes.

At 7.5 months, infants' abilities to recognize familiarized words (i.e., recognition scores) were correlated with productive vocabulary percentile at 24 months for both simple and complex tasks. Furthermore, the Bayley MDI correlated significantly with the simple segmentation task but not the complex task yet the complex segmentation task was more highly correlated with vocabulary outcomes than the simple segmentation task. It is possible that the task demands of the Bayley Scales overlapped to a greater extent with those of the simple segmentation task than the complex task. Potentially, a normed assessment that demanded more similarity-based categorization prior to responding would have correlated with the complex segmentation task.

Two outcome groups were established based on where individuals were positioned relative to the median at 24 months. These groups began to differ in vocabulary size at 12 months and were associated with different growth trajectories over the testing period. In particular, the high outcome children's growth trajectories were better characterized by higher order, non-linear growth models and the low outcome children's trajectories more closely fit lower order models. Although both groups succeeded at the simple segmentation task and at recognition of the matched item on the complex task, only the high vocabulary group showed the ability to generalize across mismatched forms, accounting for a significant difference in recognition scores between groups for the complex segmentation task. This invites a revision of previous conclusions stating that infants at 7 to 8 months cannot recognize mismatched forms (Bortfeld & Morgan, 2011, Houston & Jusczyk, 2000; Singh, et al., 2004; 2008a; 2008b) as infants with prospects for high vocabulary at 24 months are able to succeed at a generalization task as early as 7.5 months. It is therefore possible that precocity in this aspect of early word segmentation may stand a learner in good stead for word-to-meaning mapping although obviously a causal account of this relationship cannot be confirmed by the current study.

These results present the first prospective evidence of continuity between infant word segmentation tasks and later vocabulary development. However, there are several further directions invited by these findings. First, the most intriguing question is what the basis of this continuity may be. Infant word segmentation and child lexical development likely involve many common factors (e.g. forming associations, retention of linguistic and non-linguistic detail, information processing, sensory acuity, categorical perception). It is therefore likely that the basis of any continuity is multifactorial, but the specific factors as well as the relative loading of the relevant factors remains to be discovered. In the current study, we have attempted to explore the role of limited additional factors such as the

efficiency of information processing and total attention to task. These results demonstrated that other extra-linguistic measures, specifically the number of trials required to attain familiarization, correlated negatively with outcome vocabulary size, suggesting that infants who were 'fast familiarizers' were associated with higher values in later vocabulary development. However, it is important to note that these findings measure familiarization to *linguistic* stimuli specifically, which may not generalize to other forms of information processing. In a longitudinal study designed to address the linguistic specificity of infant predictor-outcome relationships, Rose, Feldman and Jankowksi (2009) did not find evidence of predictive relationships between attention to visual recognition memory tasks and habituation to visual stimuli with later language outcomes, arguing against cross-domain attentional predictors of later language. By contrast, Rose and her colleagues found that accuracy of performance in visual recognition memory tasks (e.g., the Fagan task) and representational competence tasks (e.g., cross-modal transfer) administered at 12 months did predict later vocabulary development at 36 months. This finding suggests a domain-general relationship between performance accuracy in infant visual memory tasks and later language outcomes. The issue of domain-generality was also investigated by Kuhl and her colleagues (Kuhl, Conboy, Padden, Nelson & Pruitt, 2005) demonstrating that language measures at 24 months were significantly associated with improved discrimination of a native language phoneme contrast and reduced accuracy with a non-native phoneme pair. This suggests that general auditory sensitivities may not be wholly responsible for the continuity observed but rather that acquisition of native language structure may underlie growth in infancy and childhood. Therefore, there is a mixed account of domain-generality versus specificity with respect to the predictors of language outcomes, but currently, there are very few investigations in this domain, thus making it difficult to conclude whether predictors of child language are domain- or even task-specific. However, there is a rich literature on emerging numerical, conceptual and object knowledge in infancy all requiring perceptual categorization mechanisms analogous to those involved in infant speech processing tasks (e.g., Cohen & Younger, 1983; French, Mareschal, Mermillod, & Quinn, 2004; Quinn, Eimas & Rosencrantz., 1993, Bloom & Wynn, 1997; Wynn, 1995). These tasks could be usefully incorporated into a longitudinal prospective design, in combination with infant speech processing tasks, to determine the relative predictive power of measures associated with different cognitive domains.

In addition to considering the specificity of infant variables, it is equally important to consider the stability of our outcome measures. At 24 months, vocabulary is an area of tremendous growth and high variability, enabling promising exploration of individual differences. However, the present study did not examine mastery of phonological systems nor did it examine morphosyntactic development or the other traditional areas of linguistic analysis equally crucial to language development. It remains to be seen whether predictor variables such as word segmentation relate to outcome measures outside of the lexicon. Certainly, there is considerable evidence to suggest a relationship between vocabulary size and early grammatical development in typical learners (Bates, Bretherton & Synder, 1998; Marchman & Bates, 1994) and in atypical learners (Moyle, Weismer, Evans & Lindblom, 2007). This interdependence of grammatical and lexical knowledge forms the basis of the critical-mass hypothesis (Marchman & Bates, 1994) that presupposes the emergence of grammar once vocabulary size necessitates syntactic elaboration. On this view, one would hypothesize performance in related but later-developing domains, such as morphosyntactic suppliance, to correlate with infant word segmentation scores. Outcomes obtained at later points could address this issue more directly.

The search for clinically relevant predictors in infancy is a natural application of the individual differences approach to language acquisition. Each year, a significant contingent of preschool children (3-7%) appears to not develop language on course (Rescorla, 1989).

The vast majority of children diagnosed with a language delay or disorder are identified as having Expressive Language Delay (ELD). Fortunately, the prognosis for ELD is often quite positive, and affected children benefit greatly from therapy and remediation (Rescorla, 1989), and some children are often able to reach age-appropriate levels of language and education with successful therapeutic efforts. By contrast, those children who do not receive treatment in a timely fashion are likely to continue to be delayed through the elementary and middle school years and beyond (Rescorla & Schwartz, 1990). A widely attested finding is that children treated prior to the age of 3 are more likely to 'catch up' to their peers compared with those treated after the age of 3 (Dale, Price, Bishop & Plomin, 2003; Rescorla, Roberts & Dahlsgaard, 1997). If the variability in childhood vocabulary development could be traced to variability in tasks assessing word knowledge in infants, this would lend considerable credibility to a search for early risk markers of expressive language delay in preverbal infants. As indicated in the current set of findings, low vocabulary outcome children appear to be able to segment words that do not differ in surface form at 7.5 months similar to their high vocabulary outcome peers, but the groups appear to differ in two ways. First, low vocabulary outcome children showed lower recognition scores than their high vocabulary peers in complex segmentation tasks and marginally so even in the simple segmentation task. Perhaps a more qualitative distinction is drawn by recognition of mismatched forms whereby low vocabulary outcome children appear to be unable to generalize across variable instances of words at 7-8 months, whereas high vocabulary infants are able to do so.

In conclusion, this study confirms that individual differences in performance on infant word segmentation can reflect meaningful distinctions in later vocabulary development. Although research in infant speech processing has attributed an impressive fund of language processing capacities to young learners, the link between these abilities and later language development is just beginning to be elucidated. At each level of the language code, there is now compelling evidence linking infant capabilities with later language acquisition, which lends great promise to the utility of studying individual differences as a means of predicting individual pathways to language. The current findings contribute to this emergent field of study by demonstrating continuity in the domain of word learning between infancy and early childhood.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

REFERENCES

Aslin, RN.; Woodward, JZ.; LaMendola, NP.; Bever, TG. Models of word segmentation in fluent maternal speech to infants. In: Morgan, JL.; Demuth, K., editors. Signal to syntax: Bootstrapping from speech to grammar in early acquisition. Lawrence Erlbaum Associates; Mahwah, NJ: 1996. p. 117-134.

Bates, E.; Bretherton, I.; Snyder, L. From first words to grammar: Individual differences and dissociable mechanisms. Cambridge University Press; New York: 1988.

Bayley, N. Bayley Scales of Infant Development. 2nd Ed.. The Psychological Corporation; San Antonio, TX: 1993.

Benasich AA, Tallal P. Infant discrimination of rapid auditory cues predicts later language impairment. Behavioral Brain Research. 2002; 136(1):31–49.

Bloom P, Wynn K. Linguistic cues in the acquisition of number words. Journal of Child Language. 1997; 24:511–533. [PubMed: 9519584]

Bortfeld H, Morgan J. Is early word-form processing stress-full? How natural variability supports recognition. Cognitive Psychology. 2010; 60:241–266. [PubMed: 20159653]

Bortfeld H, Morgan JL, Golinkoff R, Rathbun K. Mommy and Me. Familiar names help launch babies into speech-stream segmentation. Psychological Science. 2005; 16:298–304. [PubMed: 15828977]

Brent, Michael. An efficient, probabilistically sound algorithm for segmentation and word discovery. Machine Learning. 1999; 34:71–105.

Cohen, LB.; Younger, BA. Perceptual categorization in the infant. In: Scholnick, E., editor. New trends in conceptual representation. Erlbaum; Hillsdale, NJ: 1983. p. 197-220.

Cristia, A.; Seidl, A. Sensitivity to prosody at 6 months predicts vocabulary at 24 months. In: Danis, N.; Mesh, K.; Sung, H., editors. BUCLD 35: Proceedings of the 35th annual Boston University Conference on Language Development; Somerville: Cascadilla Press. 2011;

Dale PS, Price TS, Bishop DVM, Plomin R. Outcomes of early language delay: I. Predicting persistent and transient language difficulties at 3 and 4 years. Journal of Speech, Language, & Hearing Research. 2003; 46:544–560.

Dale PS, Bates E, Reznick JS, Morisset C. The validity of a parent report instrument of child language at twenty months. Journal of Child Language. 1989; 16:239–249. [PubMed: 2760125]

Fenson, L.; Dale, P.; Reznick, S.; Thal, D.; Bates, E.; Hartung, J.; Pethick, S.; Reilly, J. MacArthur-Bates Communicative Development Inventories. Singular; San Diego: 1993.

Fernald A, Perfors A, Marchman V. Picking Up Speed in Understanding: Speech Processing Efficiency and Vocabulary Growth Across the 2nd Year. Developmental Psychology. 2006; 42(1): 98–116. [PubMed: 16420121]

French RM, Mareschal D, Mermillod M, Quinn PC. The role of bottom-up processing in perceptual categorization by 3- to 4-month-old infants: simulations and data. Journal of Experimental Psychology: General. 2004; 133(3):382–397. [PubMed: 15355145]

Gibbs, ED.; Teti, DM. Interdisciplinary assessment of infants: A guide for early intervention professionals. Paul H. Brookes; Baltimore: 1990.

Houston DM, Jusczyk PW. The role of talker-specific information in word segmentation by infants. Journal of Experimental Psychology: Human Perception and Performance. 2000; 26:1570–1582. [PubMed: 11039485]

Houston DM, Horn DL, Qi R, Ting JY, Gao S. Assessing speech discrimination in individual infants. Infancy. 2007; 12(2):119–145.

Jusczyk PW, Aslin RN. Infants' detection of the sound patterns in fluent speech. Cognitive Psychology. 1995; 28:1–23. [PubMed: 7895467]

Jusczyk, PW. The Discovery of Spoken Language. MIT Press; Cambridge: 1997.

Jusczyk PW, Houston D, Newsome M. The beginnings of word segmentation in English-learning infants. Cognitive Psychology. 1999; 39:159–207. [PubMed: 10631011]

Kemler Nelson DG, Jusczyk PW, Mandel DR, Myers J, Turk A, Gerken L. The headturn preference procedure for testing auditory perception. Infant Behavior and Development. 1995; 18:111–116.

Kuhl PK. Early language acquisition: cracking the speech code. Nature Reviews Neuroscience. 2004; 5:831–843.

Kuhl PK, Conboy BT, Padden D, Nelson T, Pruitt J. Early speech perception and later language development: implications for the "critical period.". Language Learning and Development. 2005; 1(3&4):237–264.

Ma W, Golinkoff R, Houston DM, Hirsh-Pasek K. Word learning in infant- and adult-directed speech. Language Learning and Development. 2011; 7(3):185–201.

Mandel DR, Jusczyk PW, Pisoni DB. Infants' recognition of the sound patterns of their own names. Psychological Science. 1995; 6:314–317.

Marchman V, Bates E. Continuity in lexical and morphological development: A test of the critical mass hypothesis. Journal of Child Language. 1994; 21:339–366. [PubMed: 7929685]

Marchman VA, Fernald A. Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. Developmental Science. 2008; 11(3):F9–16. [PubMed: 18466367]

Mintz TH. Frequent frames as a cue for grammatical categories in child directed speech. Cognition. 2003; 90(1):91–117. [PubMed: 14597271]

Molfese DL, Molfese VJ. Electrophysiological indices of auditory discrimination in newborn infants: The bases for predicting later language development. Infant Behavior and Development. 1985; 8:197–211.

Molfese DL, Molfese VJ. Discrimination of language skills at five years of age using event related potentials recorded at birth. Developmental Neuropsychology. 1997; 13:135–156.

Molfese VJ, Molfese DL. Environmental and social influences on reading skills as indexed by brain and behavioral responses. Annals of Dyslexia. 2002; 52:121–137.

Molfese, VJ.; Molfese, DL.; Molfese, PJ. Relation between early measures of brain responses to language stimuli and childhood performance on language and language-related tasks. In: Donna, Coch; Geraldine, Dawson; Kurt, Fischer, editors. Human behavior, learning, and the developing brain: Atypical development. Guilford; New York: 2007.

Moyle MJ, Weismer SE, Evans JL, Lindstrom MJ. Longitudinal relationships between lexical and grammatical development in typical and late-talking children. Journal of Speech, Language, and Hearing Research. 2007; 50:508–528.

Newman RS, Bernstein Ratner N, Jusczyk AM, Jusczyk PW, Dow KA. Infants' early ability to segment the conversational speech signal predicts later language development: a retrospective analysis. Developmental Psychology. 2006; 42(4):643–655. [PubMed: 16802897]

Quinn PD, Eimas PD, Rosencrantz SL. Evidence for representations of perceptually similar natural categories by 3-month-old and 4-month-old infants. Perception. 1993; 22:463–475. [PubMed: 8378134]

Rescorla L. The Language Development Survey: a screening tool for delayed language in toddlers. Journal of Speech and Hearing Disorders. 1989; 54:587–599. [PubMed: 2811339]

Rescorla L. Age 9 outcomes for late-talking toddlers. Journal of Speech, Language, and Hearing Research. 2002; 45:360–371.

Rescorla L. Age 17 language and reading outcomes in late-talking toddlers: support for a dimensional perspective on language delay. Journal of Speech, Language, and Hearing Research. 2009; 52:16–30.

Rescorla, L. Rating scale systems for assessing psychopathology: The Achenbach System of Empirically Based Assessment (ASEBA) and the Behavior Assessment System for Children-2 (BASC-2). Chapter. In: Matson, JM.; Andrasik, F., editors. Assessing childhood psychopathology and developmental disabilities. Springer; New York: 2008.

Rescorla L, Alley A. Validation of the Language Development Survey (LDS): A parent report for identifying language delay in toddlers. Journal of Speech, Language, and Hearing Research. 2001; 44:434–445.

Rescorla L, Roberts J, Dahlsgaard K. Late talkers at 2: Outcome at age 3. Journal of Speech and Hearing Research. 1997; 40:556–566.

Rescorla L, Schwartz E. Outcome of toddlers with specific expressive language delay. Applied Psycholinguistics. 1990; 11:393–408.

Rose SA, Feldman JF, Jankowski JJ. A Cognitive approach to the development of early language. Child Development. 2009; 80(1):134–150. [PubMed: 19236397]

Shi R, Gauthier B. Recognition of function words in 8-month-old French-learning infants. Journal of the Acoustical Society of America. 2005; 117:2426.

Singh L, Morgan JL, White KS. Preference and processing: The role of speech affect in early spoken word recognition. Journal of Memory and Language. 2004; 51:173–189.

Singh L, White K, Morgan JL. Effects of pitch and amplitude variation on infant word recognition. Language Learning and Development. 2008a; 4(2):157–178.

Singh L, Nestor SS, Bortfeld H. Overcoming Effects of Variation on Infant Word Recognition: Influences on Word Familiarity. Infancy. 2008b; 13(1):57–74. [PubMed: 21088702]

Swingley D. Lexical exposure and word-form encoding in 1.5-year-olds. Developmental Psychology. 2007; 43:454–464. [PubMed: 17352552]

Swingley D. Contributions of infant word learning to language development. Philosophical Transactions of the Royal Society B. 2009; 364:3617–3622.

Trehub SE, Henderson JL. Temporal resolution in infancy and subsequent language development. Journal of Speech and Hearing Research. 1996; 39:1315–20. [PubMed: 8959616]

Tsao F-M, Liu H-M, Kuhl PK. Speech perception in infancy predicts language development in the second year of life: A longitudinal study. Child Development. 2004; 75:1067–1084. [PubMed: 15260865]

Weber C, Hahne A, Friedrich M, Friederici AD. Reduced stress pattern discrimination in 5-month-olds as a marker of risk for later language impairment: Neurophysiologial evidence. Cognitive Brain Research. 2005; 25(1):180–187. [PubMed: 15951161]

Whatley, JL. Bayley Scales of Infant development. In: Keyser, DJ.; Sweetland, RC., editors. Test critiques. Vol. 6. Test Corporation of America; Kansas City, MO: 1987. p. 38-47.

Wynn K. Infants possess a system of numerical knowledge. Current Directions in Psychological Science. 1995; 4:172–177.

Zwaigenbaum L, Bryson S, Lord C, Rogers S, Carter A, Carver L, et al. Clinical assessment and management of toddlers with suspected autism spectrum disorder: Insights from studies of high-risk infants. Pediatrics. 2009; 123(5):1383–1391. [PubMed: 19403506]
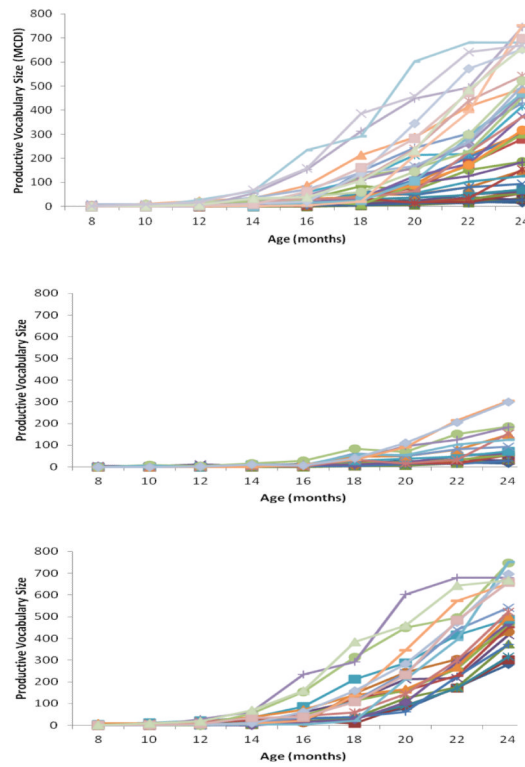
Figure 1.
a Productive Vocabulary Size as measured by the MCDI from 8 to 24 months.
b Productive Vocabulary Size as measured by the MCDI from 8 to 24 months for Infants below the Median at 24 months.
c Productive Vocabulary Size as measured by the MCDI from 8 to 24 months for Infants above the Median at 24 months.
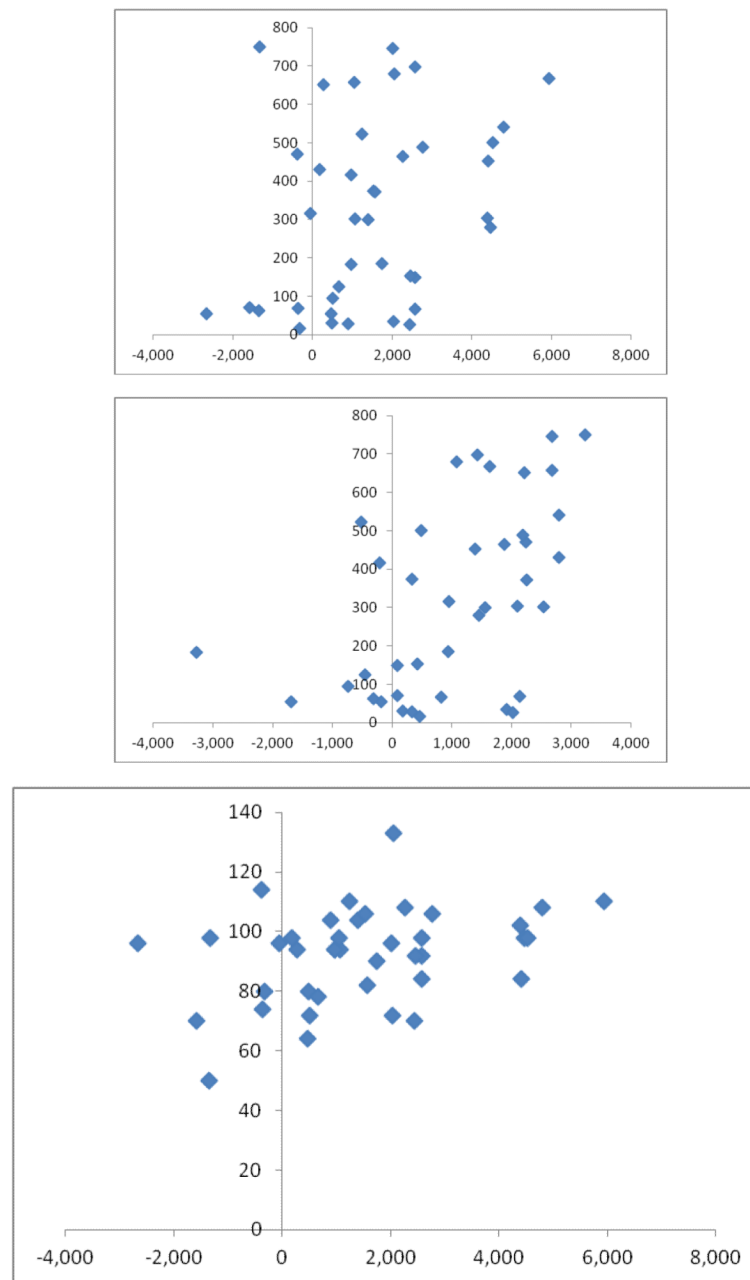
Figure 2.
a Scatterplot of recognition scores for simple segmentation task and vocabulary size at 24 months (y-axis: MCDI 23 months; x-axis: recognition scores).
b Scatterplot of recognition scores for complex segmentation task (averaged across items) and vocabulary size at 24 months (y-axis: MCDI at 23 months; x-axis: recognition scores).
c Scatterplot of recognition scores for simple segmentation task and Bayley MDI at 23 months (y-axis: Bayley MDI at 23 months; x-axis: recognition scores).
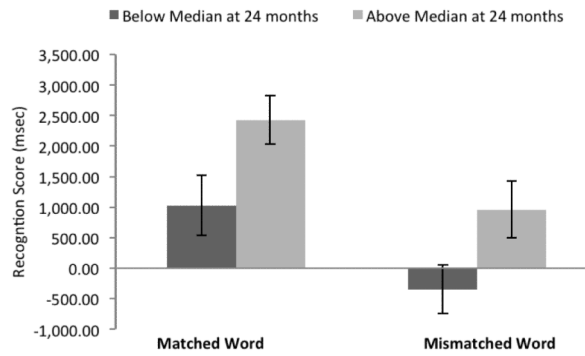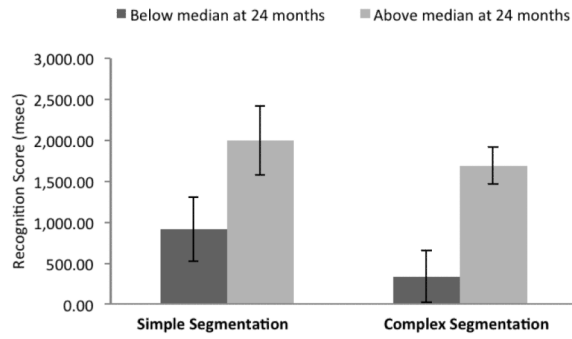
Figure 3.
a Recognition scores for simple and complex segmentation tasks by group (means and SEM).
b Recognitions scores for matched and mismatched items (complex segmentation task) by group (means and SEM).

Table 1

Infant Measures (Means and SD)

Simple Segmentation Recognition scores (msec) (SD)	Complex Segmentation - Mean scores across matched and mismatched words (SD)	Complex Segmentation - Pitch-matched Words Recognition scores (msec) (SD)	Complex Segmentation - Pitch-mismatched Words Recognition scores (msec) (SD)	Overall Listening Times during Recognition Passages (simple and complex averaged) (SD)	Number of Familiarization Trials to Criterion (simple and complex averaged) (SD)
1,482.48 (1881.15)	1048.24 (1372.01)	1,759.84 (2076.36)	336.66 (2038.31)	33451 (7905.64)	3.05 (.58)

Table 2

Outcome Measures (Means and SD)

MCDI (mean productive vocabulary size and SD)									Bayley MDI (mean standard score and SD)
8 months	10 months	12 months	14 months	16 months	18 months	20 months	22 months	24 months	23 months
0.83	2.03	5.18	14.53	33.73	78.93	137.25	217.83	319.68	92.28
1.88	3.19	7.23	17.01	48.37	89.48	139.97	187.19	239.41	15.80

Table 3

Correlations between Infant Measures and Outcome Measures

	Simple Segmentation recognition scores	Complex Segmentation recognition scores	Mean Listening Times	No. of Familiarization Trials	24 month Vocabulary Size	Bayley MDI
Simple Segmentation recognition scores		.27	−.06	−.29	.32*	.34*
Complex Segmentation recognition scores			0.05	−0.12	.51**	.24
Mean Listening Times				−.05	−.17	−.09
No. of Familiarization Trials					−.37*	−.26
24 month Vocabulary Size						.67**
Bayley MDI						

*
Correlation is significant at the 0.05 level (2-tailed).

**
Correlation is significant at the 0.01 level (2-tailed).

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Table 4

Model Selection for Entire Sample and Each Outcome Group. Low PR indicates a poor fit to the model.

		Linear Model	Quadratic Model	Cubic Model	Quartic Model
Entire Sample	GFI	.50	.77	.80	.89
	Chi-Square	136.26 33	60.73 30	49.37	36.6
	Chi-Square DF			26	21
	PR>Chi-Square	<.0001	.0008	.0037	.0187
Low Vocabulary Outcome Group	GFI	.66	.69	-	-
	Chi-Square	43.01	38.83	-	-
	Chi-Square DF	33	30	-	-
	PR>Chi-Square	.11	.13	-	-
High Vocabulary Outcome Group	GFI	.50	.74	.76	.81
	Chi-Square	76.87	39.58	35.45	28.74
	Chi-Square DF	33	30	26	21
	PR>Chi-Square	<.0001	.11	.1	.12