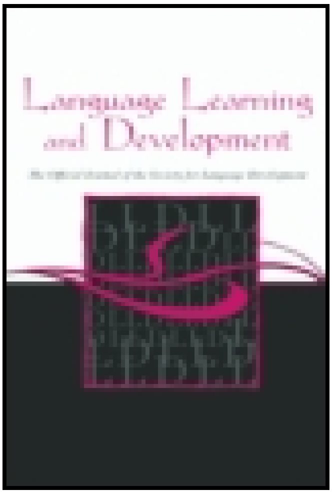
This article was downloaded by: [. Jan Edwards]
On: 14 November 2014, At: 08:17
Publisher: Routledge
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Language Learning and Development
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hlld20

Effects of Vocabulary Size on Online Lexical Processing by Preschoolers

Franzo Law II[a] & Jan R. Edwards[b]

[a] Departments of Psychology and Communication Sciences and Disorders, University of Wisconsin-Madison

[b] Department of Communication Sciences and Disorders, University of Wisconsin-Madison

Published online: 11 Nov 2014.

To cite this article: Franzo Law II & Jan R. Edwards (2014): Effects of Vocabulary Size on Online Lexical Processing by Preschoolers, Language Learning and Development, DOI: 10.1080/15475441.2014.961066

To link to this article: http://dx.doi.org/10.1080/15475441.2014.961066

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions

Language Learning and Development, 00: 1–25, 2014
Copyright © Taylor & Francis Group, LLC
ISSN: 1547-5441 print / 1547-3341 online
DOI: 10.1080/15475441.2014.961066

Routledge
Taylor & Francis Group

Effects of Vocabulary Size on Online Lexical Processing by Preschoolers

Franzo Law II
Departments of Psychology and Communication Sciences and Disorders, University of Wisconsin–Madison

Jan R. Edwards
Department of Communication Sciences and Disorders, University of Wisconsin–Madison

This study was designed to investigate the relationship between vocabulary size and the speed and accuracy of lexical processing in preschoolers between the ages of 30 and 46 months using an automatic eye tracking task based on the looking-while-listening paradigm (Fernald, Zangl, Portillo, & Marchman, 2008) and mispronunciation paradigm (White & Morgan, 2008). Children's eye gaze patterns were tracked while they looked at two pictures (one familiar object, one unfamiliar object) on a computer screen and simultaneously heard one of three kinds of auditory stimuli: correct pronunciations of the familiar object's name, one-feature mispronunciations of the familiar object's name, or a nonword. The results showed that children with larger expressive vocabularies, relative to children with smaller expressive vocabularies, were more likely to look to a familiar object upon hearing a correct pronunciation and to an unfamiliar object upon hearing a novel word. Results also showed that children with larger expressive vocabularies were more sensitive to mispronunciations; they were more likely to look toward the unfamiliar object rather than the familiar object upon hearing a one-feature mispronunciation of a familiar object-name. These results suggest that children with smaller vocabularies, relative to their larger-vocabulary age peers, are at a disadvantage for learning new words, as well as for processing familiar words.

INTRODUCTION

The ability to recognize a spoken word quickly and accurately is an integral part of language learning. Most children begin to recognize words in the first year of life and to produce words around their first birthday (e.g., Benedict, 2008). Early lexical development involves two related processes: acquiring new words (i.e., word learning) and recalling these words in meaningful communicative contexts (i.e., lexical access). Word learning involves associating sequences of

---

Correspondence should be addressed to Franzo Law II, Departments of Psychology and Communication Sciences and Disorders, University of Wisconsin–Madison, 1500 Highland Ave., Rm. 489, Madison, WI 53705. E-mail: flaw@wisc.edu phonological forms with semantic meaning and then storing them in the mental lexicon, whereas lexical access involves quickly and reliably utilizing these associations. These two processes are highly interrelated and both must be employed by the child before one can say that the child "knows" a word; a word must be *both* stored in the mental lexicon and accessed during communication. Although the earliest research on lexical development in young children focused more on word learning than on lexical access (e.g., Diesendruck, Gelman, & Lebowitz, 1998; Hall, 1991; Heibeck & Markman, 1987; Moore, Angelopoulos, & Bennett, 1999; Smith, 1999; Waxman & Booth, 2003; see for review Waxman & Lidz, 2006), some studies have explored the relationship between lexical access and vocabulary size (e.g., Charles-Luce & Luce, 2009; Fernald, Perfors, & Marchman, 2006; Walley, 1993). The present study further investigated this relationship between vocabulary size and lexical processing patterns of preschoolers between the ages of 30 and 46 months.

Online Lexical Processing in Children

Much recent research has focused on lexical access—how young children quickly and reliably recognize familiar words. One widely used experimental paradigm for this research is the looking-while-listening (LWL) paradigm (e.g., Fernald et al., 2008; Marchman & Fernald, 2008), an adaptation of the inter-modal preferential looking paradigm (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987). In this paradigm, pictures of two familiar objects are presented on a computer screen and one of the two object-names is presented aurally (e.g., *find the doggie*). Reaction time (*latency*) of looking to the target is quantified as the time from target word onset to when a child first looks to the target image (measured only on trials on which the child was looking at the distractor picture at target word onset); *accuracy* is quantified as the number of looks to the target image relative to the total number of looks to either the target or distractor within a specified time window. Using the LWL paradigm in a longitudinal study, Fernald et al. (2006) found that both latency and accuracy of looking to highly familiar words improved systematically from 15 to 25 months. That is, during a time of rapid expansion in vocabulary size (Benedict, 2008; Goldfield & Reznick, 2009), children also become faster and more accurate at recognizing familiar words. An analysis of individual differences found that vocabulary size (both receptive and expressive, as measured by the Macarthur Bates Communicative Development Inventory (MCDI; Fenson et al., 1993) at the ages of 12, 18, and 21 months predicted both latency and accuracy of looking to familiar words at 25 months. In a study that followed children from 18 to 24 months from both low- and middle-socioeconomic status (SES) families, Fernald and colleagues found similar results as in the earlier study (Fernald, Marchman, & Weisleder, 2013). In this study, children from low-SES families had generally smaller vocabulary sizes than their peers from middle-SES families, a finding consistent with previous research (Hart & Risley, 1995; Hoff, 2003). Children from low-SES families also showed less efficient lexical processing than their age peers from middle-SES families, as evidenced by lower accuracy levels and longer latencies at both 18 and 24 months. Fernald and colleagues interpreted this result to support the claim that online lexical processing is linked to vocabulary size, regardless of whether differences in vocabulary size are related to endogenous child-internal factors such as individual differences in attention (McCall & Carriger, 1993), or to environmental factors such as the quantity and quality of linguistic input to the child.

Relationship Between Vocabulary Size and Lexical Processing

Speech perception models have suggested that increases in vocabulary size allow for more efficient, phonologized perceptual routines for word recognition (Curtin & Werker, 2007; Strange, 2011; Werker & Curtin, 2005). This has borne out empirically; Fernald, Swingley, and Pinto (2001) found that 18- to 21-month-old infants with larger vocabularies were faster and more accurate at identifying familiar two-syllable words when only the first syllable was presented compared with age-matched peers with smaller vocabularies. These results suggest that infants with larger vocabularies were able to use sublexical information for lexical retrieval. The ability to identify a word with only partial acoustic information is a necessary skill for rapid lexical processing that has been consistently found in studies of lexical access of adult populations (e.g., Allopenna, Magnuson, & Tanenhaus, 1998; Grosjean, 1980; Marslen-Wilson & Warren, 1994; Smits, Warner, McQueen, & Cutler, 2003). Several infant studies also support the claim that larger vocabularies are associated with more finely detailed phonological organization in the lexicon. For instance, Werker, Fennell, Corcoran, and Stager (2002) found that 18-month-old children with larger expressive and receptive vocabularies were better than age-matched peers with smaller vocabularies at differentiating novel words that differed by only a single distinctive feature. Similarly, Graf Estes, Edwards, and Saffran (2011) found that 18-month-olds with larger receptive vocabularies showed a larger effect of phonotactic probability on word learning than age peers with smaller vocabularies. The effect of vocabulary size in language processing has been demonstrated within-subject, as well; bilingual children are faster and more accurate in identifying words of the language in which they have a larger vocabulary and more relative linguistic exposure (Hurtado, Grüter, Marchman, & Fernald, 2014). A similar relationship between vocabulary size and semantic specificity has also been found in children with relatively smaller vocabulary sizes, such as late talkers (Beckage, Smith, & Hills, 2010; Colunga & Sims, 2011, 2012; Jones & Smith, 2005) and children with specific language impairment (Sheng & McGregor, 2010). Thus, it may be that a large vocabulary leads to a better-organized lexicon, which in turn results in more efficient lexical processing.

Furthermore, it is likely that children with larger vocabularies have more robust lexical knowledge, relative to children with smaller vocabularies, even of words that are familiar to both groups of children. Recent research suggests that, for both children and adults, knowledge of new words develops gradually, rather than in a single fast-mapping context (McMurray, 2007; Smith & Yu, 2008; Vouloumanos, 2008). For example, Smith and Yu found that 12- and 14-month-olds gradually built up phonological/semantic paired associations for novel words based on statistical learning. If children gradually develop more robust lexical representations based on linguistic experience, then one would expect children with larger vocabularies (i.e., children who produce and comprehend more words in more contexts) to have more efficient lexical processing precisely because of these more highly developed lexical representations.

Lexical Processing and Encounters with Novel Words

Lexical processing involves more than simply processing familiar words; lexical access models have described novel word perception as the process of accessing the lexicon of known words, comparing these words to the novel word in question, and determining that this word is not yet known (e.g., Luce, Goldinger, Auer, & Vitevitch, 2000; Marslen-Wilson, 1987; Marslen-Wilson & Welsh, 1978; Vitevitch & Luce, 1998). A large body of research has shown that if children hear a novel word when presented with an unfamiliar object and one or more familiar objects, they will associate the novel word with the unfamiliar object–a process known as disambiguation (e.g., Golinkoff, Hirsh-Pasek, Bailey, & Wenger, 1992; Markman, 1991; Merriman, Bowman, & MacWhinney, 1989). There is some debate regarding the underlying mechanism that drives disambiguation. It is possible that disambiguation is due to the child not having a label for the unfamiliar object, and this lexical gap drives the association of the novel object to the unfamiliar object (Novel Name-Nameless Category (N3C) Principle: Mervis & Bertrand, 1994). However, at least some aspect of disambiguation is evident before the child has substantial language skills. Infants have demonstrated patterns of disambiguation and associations between novel object and unfamiliar object well before they have amassed a large vocabulary (Dewar & Xu, 2007; Markman, Wasow, & Hansen, 2003; Mather & Plunkett, 2010). This tendency may be independent of vocabulary size, but rather is guided by a preference for having only one label for an object (i.e., mutual exclusivity: Markman, 1990; Markman et al., 2003), or by the social-pragmatic assumption that the speaker is likely referring to the novel object when using the novel word (Diesendruck & Markson, 2001).

Although typically developing infants readily demonstrate disambiguation before acquiring a large vocabulary, a recent study by Bion and colleagues suggests that this ability becomes more efficient with age and that children with larger vocabularies are better able to capitalize on disambiguation in order to acquire new words. This study investigated lexical processing in 18-, 24-, and 30-month-olds using the LWL paradigm, but with a familiar and an unfamiliar object presented during each trial, and either a familiar word or a novel word presented aurally (Bion, Borovsky, & Fernald, 2013). They found a relationship between accuracy and age; the 24- and 30-month-olds, but not the 18-month-olds, were significantly more likely to look at the unfamiliar object upon hearing the novel word. This finding is similar to that of Mather and Plunkett (2009, 2010), who found that 22.5-month-olds, but not 19.5-month-olds, demonstrated disambiguation upon the first exposure to the novel word and unfamiliar object; younger children require multiple presentations of the novel word and object to demonstrate disambiguation. Bion et al. (2013) also observed that accuracy of looking to the unfamiliar object when hearing a novel word was associated with vocabulary size for the 24- and 30-month-olds. Children with larger vocabularies were even more likely to look to the unfamiliar object when a novel word was presented.

Lexical Processing and Differentiating Novel Words from Mispronunciations of Known Words

Although children readily assign novel words to novel objects, this ability is mediated by the *degree* to which the novel word differs from other words that the child knows. The more phonologically similar the novel word is to a known word, the less likely the child is to exhibit disambiguation (Mather & Plunkett, 2011). This parallels adult speed in word processing, which is inversely proportional to the number of phonologically similar words in the adult's lexicon (Luce & Pisoni, 1998; Marslen-Wilson & Welsh, 1978). Furthermore, the child must learn to differentiate between phonologically similar words and mispronunciations of known words; upon hearing a production such as *vaby* or *maybe*, the child must determine whether that production is a mispronunciation of *baby*, or a novel word. There are many opportunities for mispronunciations in children's environments. However, in addition to mispronunciations, there are also many correctly pronounced words that differ from one another by only a single phoneme, particularly in English. Children need to be able to identify whether a word is familiar or novel quickly and accurately, even if it is similar to a word they already know. Children, as well as adults, have shown sensitivity to mispronunciations of familiar words, both behaviorally (e.g., Bailey & Plunkett, 2002; Ballem & Plunkett, 2005; Mani & Plunkett, 2007; Merriman & Schuster, 1991; Swingley & Aslin, 2000, 2002; White & Morgan, 2008) and neurophysiologically (e.g., Duta, Styles, & Plunkett, 2012; Friedrich & Friederici, 2005).

Swingley and Aslin (2000, 2002) used the LWL paradigm to test sensitivity to mispronunciations of familiar words. Children were presented with pictures of two familiar objects on a computer monitor and heard either a correct production or a mispronunciation (e.g., *baby* or *vaby*) of one of the object-names. Swingley and Aslin (2000) examined responses of 18- to 23-months-olds and found that they were less likely to look to the familiar object upon hearing its label mispronounced, compared to correct productions, suggesting that children were sensitive to the mispronunciations. They also observed that the effect of mispronunciation was similar across the age range tested; furthermore, the effect of mispronunciation on accuracy and latency was not associated with vocabulary size. In Swingley and Aslin (2002), the participants were 14-month-olds and the degree of mispronunciation was manipulated by changing the number of distinctive features of the correct pronunciation that were altered in order to achieve a mispronunciation. Although the 14-month-olds were sensitive to mispronunciations of known words, they were not sensitive to the degree of mispronunciation. This finding is similar to that of Bailey and Plunkett (2002), who observed sensitivity to mispronunciations in the looking patterns of both 18- and 24-month-olds, but found no differences in looking patterns as a function of whether the mispronunciation differed by one or two distinctive features from the correct pronunciation. Bailey and Plunkett (2002), like Swingley and Aslin (2000, 2002), found no influence of vocabulary size on sensitivity to mispronunciation. Using the same paradigm, Ballem and Plunkett (2005) examined 14-month-olds' sensitivity to mispronunciations of familiar and newly learned words, whereas Mani and Plunkett (2007) examined 15-, 18-, and 24-month-olds sensitivity to both vowel and consonant mispronunciations in familiar words. Both studies found that children were sensitive to mispronunciations across the experimental conditions, but that there was no effect of vocabulary size on responses to mispronunciations. However, the mispronunciation paradigm used in all of these studies presented two familiar objects within a trial (e.g., an image of a baby and a dog); thus, the choice of where to look, upon hearing a mispronounced word (e.g., *vaby*), would likely be biased to assuming that the mispronunciation is not a novel word.

White and Morgan (2008) modified the mispronunciation task to minimize this bias. In this study, a familiar object and an unfamiliar object were presented during a trial. There were three conditions for the auditory stimuli: correct productions of familiar object names, mispronunciations of the familiar object names, and nonwords. Thus, children had a choice of looking to the familiar object or to an unfamiliar object when they heard a mispronunciation. White and Morgan found that 19-month-old infants looked at the familiar object when its object name was presented and at the unknown object when the nonword was presented. Children looked at both the familiar and unfamiliar objects when the mispronunciation was presented, with a decrease in relative looking time to the familiar object relative to the degree of mispronunciation. These results contrast with those of Swingley and Aslin (2002) and Bailey and Plunkett (2002); both of these studies found that latency and accuracy for mispronunciations were unrelated to the number of distinctive features that differentiated the correct pronunciation and the mispronunciation. The fact that White and Morgan (2008) but not Swingley and Aslin (2002) or Bailey and Plunkett (2002) found that children are sensitive to the degree of mispronunciation suggests that the paradigm used by White and Morgan provides a better measure of children's sensitivity to mispronunciations. However, White and Morgan did not examine whether sensitivity to mispronunciation varied as a function of vocabulary size or age.

The Present Study

The purpose of the current study was to continue to explore the relationship between vocabulary size and lexical processing efficiency. We chose to test children who were somewhat older than the children in most of the studies cited above. With the exception of the Bion et al. (2013) study, which included 30-month-olds, all of the other studies examined lexical processing in children between 14 and 25 months. There are large differences in vocabulary size in the second year of life, but much of this early variability disappears by the preschool years (e.g., Paul, 1993; Rescorla, Mirak, & Singh, 2000). It is possible that early differences in lexical processing efficiency are related to this early variability in vocabulary size; these differences may be smaller or even disappear in preschool-aged children. Alternatively, these differences might continue to be observed or even increase with age.

This study adapted the paradigm used by White and Morgan (2008). This allowed us to examine three language skills within in a single experimental task: (1) children's identification of familiar words, (2) their disambiguation of novel words (i.e., their tendency to look to the unfamiliar object and not the familiar object upon hearing a novel word), and (3) their sensitivity to small phonetic differences (i.e., mispronunciations). This study explored three related questions. First, is lexical processing efficiency for familiar words related to vocabulary size in preschool-aged children? Second, is processing efficiency for novel words relative to vocabulary size in the same children? Finally, is processing efficiency for mispronunciations of familiar words related to vocabulary size in this same group of children? If the relationship between vocabulary size and lexical processing efficiency that has been observed in 18- to 25-month-old children for familiar words in previous research is related to the fact that children with larger vocabularies have more finely-grained lexical organization, then we would expect that there would also be a relationship between vocabulary size and lexical processing efficiency for novel words and for mispronunciations of familiar words. In particular, we would expect that children with larger vocabularies, relative to children with smaller vocabularies, would be more likely to look to the unfamiliar object when they heard either a mispronunciation of a familiar word or a novel word.

METHODS AND PROCEDURE

Participants

The participants were 34 children (17 female, 16 male) between the age of 30 to 46 months (mean = 37.4 months, SD = 5.26). An additional 5 children were run but were excluded from the analysis (2 due to more than 50% missing data, 2 because of computer malfunction, and 1 because of inability to attend to the task). All children were from middle to upper-SES families (maternal level of education: 18 had graduate degrees, 12 had college degrees, 1 had some college, 3 declined to answer). All children had average to well above average vocabulary scores. The mean standard score on the *Expressive Vocabulary Test,* 2nd edition (EVT-2; Williams, 2007) was 128 (range = 106-149, SD = 11), and the mean standard score on the *Peabody Picture Vocabulary Test, 4th edition* (PPVT-4; Dunn & Dunn, 2007) was 129 (range = 96-159, SD = 13). All children also passed a pure-tone hearing screening (20 dB HL at 500, 1000, 2000, and 4000 Hz).

Stimuli

Six familiar words were chosen based on a number of criteria. All words were one syllable in length and had a consonant-vowel-consonant (CVC) structure. All words could easily be pictured and were familiar to at least 90% of 30-month-old children, based on data from the MCDI (*Words and Sentences, production*; Dale & Fenson, 1996). Six corresponding mispronunciations were also chosen by altering the initial consonant of the familiar words by a one-feature change. Six nonwords were generated, also having a CVC structure. Because there is some research suggesting a relationship between word learning and phonotactic probability (e.g., Storkel, 2001), the nonwords were matched to the mispronunciations of the familiar words on the basis of total phonotactic probability (CV + VC). Phonotactic probability was calculated from the Hoosier Mental Lexicon, using the procedure described in Edwards et al. (2004). See Table 1 for a list of target words used.

TABLE 1
List of the Images Used (in Italics) and the IPA of the Target Words Presented (IPA), by Condition. For correct pronunciation (CP) and mispronunciation (MP) trials, a child saw a familiar image from the CP column and an unfamiliar image from the MP column and heard either the correct pronunciation or mispronunciation of the familiar image's label. For nonword (NW) trials, a child saw an unfamiliar image from the NW column and a familiar image not included in the MP and CP trials (fourth column) and heard a nonword from the NW column

	CP	*MP*	*NW*	*Familiar Images matched with NW*
/s–ʃ/	/sup/	/ʃup/	/tʃim/	
	soup	*bamboo steamer*	*pastry mixer*	*bed*
	/ʃuz/	/suz/	/giv/	
	shoes	*chemistry flasks*	*golf club trolley*	*sock*
/d-t/	/dɑg/	/tɑg/	/veɪf/	
	dog	*wombat*	*sextant*	*ball*
	/toz/	/doz/	/fɪd/	
	toes	*concertina*	*horned melon*	*cake*
/d-g/	/dʌk/	/gʌk/	/neɪdʒ/	
	rubber duck	*rubber unidentifiable creature*	*universal work holder*	*car*
	/gɝl/	/dɝl/	/ʃæn/	
	girl	*pygmy marmoset*	*bassoon reed*	*cup*

The visual stimuli comprised two sets of 12 pairs of color real-world images. Each pair contrasted a familiar object and an unfamiliar object, matched for height (333 pixels), animacy, and complexity/interestingness, based on the authors' judgment. The familiar pictures were identified by 34 children (ages 2;2-4;0) to ensure that the pictures would be recognized as intended. All of the children who participated in the norming study were typically developing, based on teacher report, and were tested in their preschool classrooms. These children were asked to name the images, one at a time. Then each child was asked by the experimenter to point to the familiar image, presented with three other images. The selected images were labeled as expected, or were given a semantically-related label in open-set identification (e.g., "face" for *girl*, "food" for *soup*) and were all recognized by at least 32 of the 34 children in the closed-set identification. Open-set identification of the unfamiliar objects was also performed by a subset of 14 children (ages 2;2-3;0).[1] Pictures of unfamiliar objects were used if at least 75% of children either reported not recognizing the object or gave inconsistent labels (e.g., "bus," hammer," and "paintbrush" were some of the responses for *bassoon reed*).

A Tobii T60 XL eye-tracker (96 dpi) was mounted on an adjustable wall mount, so that the center of the monitor was positioned approximately 60 cm from the participant's eyes. Each of the two pictures in the pair was presented within a 600 × 600 pixel gray box (a visual angle of approximately 15 degrees). Two pictures were placed next to each other, centered on the vertical axis of the screen, 100 pixels from the screen edge and 520 pixels from each other (approximately 13 degrees).

The auditory stimuli were recorded by a young adult female native Wisconsin dialect speaker in a child-directed speech register. Reinforcer phrases such as *This is fun!* and *You're doing great!* were also recorded. Two tokens of each target item were selected. Stimuli items were equated for duration within a set of words (e.g., *dog*, its mispronunciation /tɑg/, and corresponding nonword /veɪf/) in Praat using the TD-PSOLA^{TM} algorithm (Moulines & Charpentier, 1990). All productions were produced in carrier phrases such as *Find the* ____ and *See the* ____. To ensure that there were no coarticulatory cues in the carrier phrase, neutral carrier phrases (*Find the egg.* and *See the egg.*) were also recorded. Praat was then used to remove the word *egg* and to append the target stimuli, with 80 ms silence in between the target word and carrier phrase. These sentences were normalized for average RMS amplitude and were presented at approximately 65 dB SPL.

Procedure

The experiment was designed in E-Prime® Professional 2.0 (Psychology Software Tools, Inc., 2010; Schneider, Eschman, & Zuccolotto, 2002), which was used to interface with the Tobii eye-tracker. A standard looking-while-listening paradigm was used (Fernald et al., 2008). The task was presented to the children as "watching a movie." During each trial the child was presented two images on the Tobii screen, one familiar and one unfamiliar object, with the left-right position of familiar versus unfamiliar images counterbalanced.

---

[1] Two of the children who normed the images also participated in the experiment. Because they had previous exposure with the visual stimuli, however brief, the analysis was recalculated with their data removed to ensure that including their data did not significantly alter the results.

There were three trial conditions: in the Correct Pronunciation trials (CP), the familiar object's name was presented. For the Mispronunciation trials (MP), the one-feature mispronunciation of the object name was presented. For the nonword trials, a nonword was presented. The nonword trials (NW) were presented with a different set of familiar/unfamiliar images from those used in the CP and MP trials, as shown in Table 1. These familiar images were also of CVC words known to at least 90% of 30-month-olds. Each trial began with both pictures presented in silence for 2,000 ms. After which, the carrier phrase and target word would be played. At 1,000 ms after target offset, a reinforcer phrase (such as *This is fun!*) was played. The images remained on the screen for another 1,000 ms. Trial presentation was ballistic, with a blank screen intertrial interval of 500 ms. After every six to eight trials, the child saw a still-image traverse the screen from one of the edges to the center (e.g., an image of a butterfly, or a cartoon character), paired with a reinforcer phrase, such as *You're doing great!* The reinforcer presentation was manually terminated by the experimenter. The participant's position was adjusted, or additional feedback was given when necessary during the reinforcer presentation. The reinforcer phrases and presentations also served to minimize trial-to-trial redundancy inherent in the task.

Each child was presented with two blocks of 36 trials; both blocks contained 12 CP trials, 12 MP trials, and 12 NW trials (6 words of each trial type × 2 carrier phrases). Two additional CP trials were included at the beginning of each block, in order to familiarize the child with the task; these trials used different target words and images and were excluded from analysis. The subsequent trials were pseudo-randomized within each block, so that the first test trial was a CP trial, and so that no more than two trial conditions or no more than three identical carrier phrases occurred consecutively. The pseudo-randomization also ensured that CP trials were separated by at least three trials from their corresponding mispronunciations. All children were presented the same pseudo-randomization. Stimulus pairs were yoked for both the CP/MP trials and the NW trials (e.g., *dog* and *wombat* always appeared together). A total of 72 experimental trials were presented across the two blocks, with each target word presented four times.

Each block took approximately six minutes to complete; either the hearing screening or a portion of a standardized language test was conducted between the two blocks.

STATISTICAL ANALYSIS

Most analyses of eye-tracking data have examined latency of first look to target or relative looking time to target. However, there are several limitations to these analyses. First, only about 50% of trials can be used in the latency analysis, since children must be looking at the distracter picture at word onset. Furthermore, it is difficult to determine how to use latency for analysis of the mispronunciation trials, as it is unclear whether the mispronunciation labels the familiar or the unfamiliar object. Finally, the accuracy measure is a relatively crude measure of comprehension that does not provide information regarding the changing pattern of eye gaze over time. Therefore, we decided to use a mixed-effects growth curve analysis of the eye-tracking data with the log-odds of looking to the familiar-object as the dependent variable, as proposed recently by several researchers (Barr, 2008; Mirman, Dixon, & Magnuson, 2008; Mirman, Yee, Blumstein, & Magnuson, 2011).

The starting point for the growth curve analyses was determined empirically, as proposed by Barr (2008). We plotted the grand mean of all CP trials and found that 200 ms after target word onset was the earliest time point at which the curves for the CP trials showed consistent upward movement. In addition, 200 ms is approaching the earliest time at which a child would be expected to plan and initiate an eye movement in response to the stimulus (Fernald et al., 2008; Haith, Wentworth, & Canfield, 1993). As the ending point, 1,700 ms after target word onset was chosen because a 1,500 ms window of analysis has been used in other looking-while-listening experiments with children as young as 24 months (Marchman & Fernald, 2008). Although the grand mean of looking patterns during CP trials may suggest that the maxima of looking to the familiar object was earlier than this time window, we chose to use a 1500 ms time window so as not to penalize younger children within the dataset who may take comparatively longer than the group mean to identify the familiar object reliably. In addition, using a curvilinear model captures the looking patterns of children who quickly identify the familiar image and then look away. The time window used for analysis included the presentation of the target word followed by silence and did not include the time during which the reinforcer phrase was played.

Data Reduction

Two areas of interest (AOIs) were defined by the pixel location of the gray boxes surrounding the two objects. The mean (x,y) coordinates of the left and right pupil captured by Tobii were then coded as 1 or 0, where 1 was defined as a look to the familiar object and 0 was defined as a look to the unfamiliar object. Eye gazes within the screen, but not to the AOIs were tabulated in order to quantify how often the eyes could not be tracked during the experimental trial, but were not used in analysis. Missing data due to looks away from the Tobii screen, blinking, etc., were interpolated within a 150 ms time window, provided that the recorded position of the eyes before and after the missing data point(s) were looking to the same AOI. This time window would be unlikely to include a saccade-fixation-saccade sequence where the untracked fixation was to a target other than the AOI (Inhoff & Radach, 1998; Radach, Heller, & Inhoff, 1999) and has been used in previous research investigating the eye gaze patterns of children (e.g., Wass, Smith, & Johnson, 2013).

Changes in eye tracking patterns over time within a single trial are inherently dependent, due to the physiology of eye movement. Due to the assumption of mixed-effects models that the observations of the outcome variable be independent, we used a binning procedure similar to that described in Barr (2008). Barr argued that using binning to calculate the log-odds of looking to the familiar image across multiple trials score will filter out eye-movement based dependencies of adjacent time point measures. In addition, this allowed for a more stable growth curve for each child that was less affected by missing data. We used three consecutive time points (an approximately 50 ms time window) across all trials of a given trial condition to calculate the log-odds of looking to the familiar object. This allowed for a maximum of 72 values for computing the log-odds for a given child, for a particular time bin (6 words per trial type × 2 carrier phrases × 2 blocks × 3 time points). The presence of extreme log-odds values (e.g., the log-odds is mathematically undefined if all eye tracks within a given time bin were coded as a look to the familiar object) was dealt with by adding 0.5 as an adjustment factor to the count of looks to the familiar object (numerator) and to the count of looks to the distractor (denominator) before calculating the log-odds (Mirman, 2014, p. 110).

As mentioned above, the data of children with >50% missing data across all trials were not included in analysis. After data interpolation, the subjects had approximately 20% missing data within a trial, collapsing across Condition (CP: 20% MP: 22% NW: 21%). Across 816 trials for each Condition (collapsed across participant), 19% of trials had >50% of missing data (CP: 18% MP: 19% NW: 19%). Approximately 8% of trials were completely missing (CP: 7% MP: 9% NW: 9%). There was no correlation between missing data (overall and >50%) and chronological age or expressive vocabulary size.

Statistical Model

We used the lme4 R package (Bates, Maechler, & Bolker, 2013) to calculate two-level mixed-effects growth curve models with full information maximum likelihood estimation. The first set of models compared CP trials to MP trials to detect any difference in looking patterns over time for the two trial types. The set of images used in these were identical; the only difference between these trials was whether the initial phoneme of the target word was or was not the correct pronunciation of the familiar image. The log-odds across 31 time bins were used as Level-1 observations. As in Mirman et al. (2008), we modeled linear and quadratic time as Level-1 variables (Time and Time^2, respectively) in order to estimate the slope and acceleration of the growth curve. Equation (1) represents the Level-1 regression equation:

$$Y_{t(j,k)} = \alpha_{0(j,k)} + \beta_{1(j,k)} * \text{Time}_t + \beta_{2(j,k)} * \text{Time}^2{}_t + R_{t(j,k)} \quad (1)$$

$Y_{t(j,k)}$ represents the log-odds of looking to the familiar object for a particular Child j on trials involving Condition k (e.g., CP trials) at Time Bin t. $\alpha_{0(j,k)}$ represents the average log-odds for CP trials, across all observation groups (e.g., Time Bin, Child). $\beta_{1(j,k)}$ and $\beta_{2(j,k)}$ represent the change in rate and acceleration of log-odds over time, respectively. $R_{t(j,k)}$ represents the model residual. The intercept term $\alpha_{0(j,k)}$ was allowed to be random, as well as the slopes for Time and Time^2 (i.e., these terms were expected to be different for each child and each condition).

If the time bins were analyzed, numbered as 0 to 30 for Time and 0 to 900 for Time^2, the analysis would result in a correlation between Time and Time^2 that would arise merely as a consequence of the numbering scheme. As in Mirman et al. (2008), Time and Time^2 were therefore transformed such that these variables were orthogonal to each other, effectively eliminating any correlation between the two independent variables. Another advantage of using orthogonal time was that the intercept term was not fixed to a particular time bin (i.e., Time Bin 0 would correspond to the same time bin for both Time and Time^2 using a traditional numbering scheme, but not when using orthogonal numbering); the intercept term can be interpreted as corresponding to the average shape of the growth curve.

As Mirman et al. (2008) point out, in a growth curve analysis of a between-subjects experiment, the "smallest grain of analysis is the combination of subject and condition" (p. 481). We included two non-nested random effects for all models: Participant and Participant*Condition. Time and Time^2 (i.e., the Level-1 variables) are nested within these random variables. Given that all participants received all three trial conditions, Participant and Condition are fully crossed. By including Participant in the model, it was possible to capture the variance that is associated with each participant. Participant*Condition captured the variability associated with the

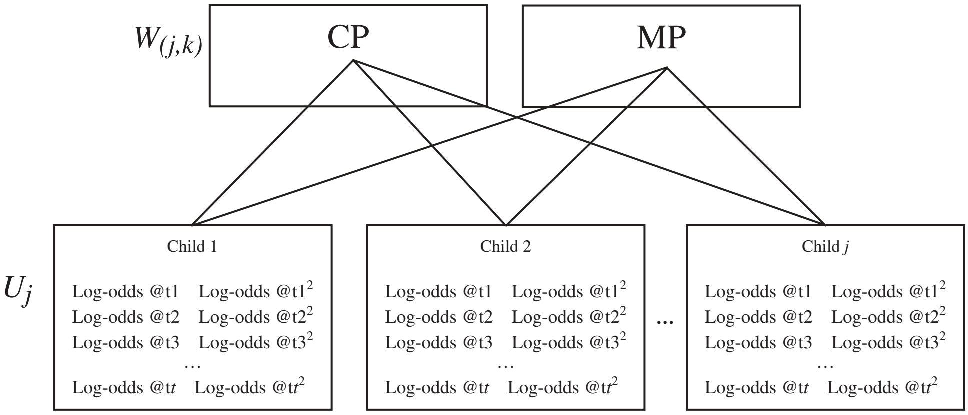


FIGURE 1 A schematic of the relationship of the random effects included in the analysis. Child (U_j) was included as a random effect to control for intersubject variability. Each child received all three trial types and thus measures of log-odds by Condition were not independent; $W_{(j,k)}$ controlled for this dependency.

dependency of the three log-odds values that was contributed by each participant for a given time bin. This second random effect is essentially an interaction between the random effect of Participant and the fixed effect of Condition. Inclusion of this interaction random effect did not change any of our interpretations of the model, but did yield a model with standard error measures that controlled for dependencies in the dataset that otherwise would have not been taken into account. The relationship between the crossed random variables is shown in Figure 1. In this figure, the bottom row of boxes represents the sets of time bins that each child contributed to each of three trial conditions, represented by the top row of boxes.

A series of mixed-effects models with varying Level-2 equations were computed in order to determine which Level-2 variables (in addition to trial condition) provided the best statistical model fit of the data. All variables were group mean-centered before analysis. First, two models were calculated with either expressive vocabulary size (EVT-2 growth scale values) or receptive vocabulary size (PPVT-4 growth scale values) as Level-2 variables, in addition to trial condition. The EVT-2 scores and PPVT-4 scores were correlated; including both variables in a single model was not feasible, due to multicollinearity. Because vocabulary growth changes exponentially over time, growth scale values (GSV) were used instead of raw scores because these scores provide a linear measure of each child's expressive vocabulary size. That is, growth scale values, in contrast to raw scores, are based on an equal-interval scale and thus are more appropriate for tracking change over time in clinical settings, as well as for use in statistical analyses (Williams, 2007).

Equations (2) to (4) represent the Level-2 regression equations, where "CL" represents either PPVT-4 GSV, EVT-2 GSV, or Age as the child-level variable included in the model:

$$\alpha_{0jk} = \gamma_{00} + \gamma_{0CL}{}^{*}\ \text{CL} + \gamma_{0MP}{}^{*}\ \text{MP} + \gamma_{0CL*MP}\left(\text{CL}^{*}\ \text{MP}\right) + \gamma_{0Age}{}^{*}\text{Age} + U_{0j} + W_{0(j,k)} + \zeta_{0jk} \tag{2}$$

$$\beta_{1jk\,\text{Time}} = \gamma_{10} + \gamma_{1CL}{}^{*}\,\text{CL} + \gamma_{1MP}{}^{*}\,\text{MP} + \gamma_{1CL*MP}\left(\text{CL}^{*}\,\text{MP}\right) + \gamma_{1Age}{}^{*}\text{Age} + U_{1j} + W_{1(j,k)} + \zeta_{1jk} \tag{3}$$

$$\beta_{2jk\,\text{Time}^2} = \gamma_{20} + \gamma_{2CL}{}^{*}\,\text{CL} + \gamma_{2MP}{}^{*}\,\text{MP} + \gamma_{2CL*MP}\left(\text{CL}^{*}\,\text{MP}\right) + \gamma_{2Age}{}^{*}\text{Age} + U_{2j} + W_{2(j,k)} + \zeta_{2jk} \tag{4}$$

Since CP trials were coded as the reference condition, γ_{00} in the above equations represent the average intercept on CP trials across all children. γ_{10} is the average slope for Time, γ_{20} for Time^2 on CP trials. The other two equations represent the effects that the Level-2 have on the parameter estimates for Time and Time^2. It is useful to think of these estimates as interactions among the variables. The random effects of Participant (U_j) and Participant*Condition ($W_{(j,k)}$) are also included in these equations. ζ_{0jk}, ζ_{1jk}, and ζ_{2jk} represent the variation of intercept and slopes for Time and Time^2 for Child j on Condition k.

Finally, mixed-effects models including a measure of vocabulary size (either EVT-2 or PPVT-4), Age, and the interaction between the two as Level-2 variables were calculated. Because there is a high correlation between chronological age and growth scale values (a measure that is expected to increase with age), standard scores of vocabulary were used in these models. The Level-1 equation for this model is identical to Equation 1. The Level-2 equations are similar to Equations (2) to (4), but with three main effects instead of two (e.g., Age, EVT-2 standard score, and Condition), and all interactions among the main effects.

In the same fashion, a set of mixed-effect growth models were calculated separately for the NW trials. The NW trials were analyzed separately from the CP and MP trials because these trials were presented with a different set of pictures from the CP and MP trials and, as noted above, the familiar image was never named during nonword trials. The Level-1 equation was the same as described above and is represented by Equation (1). The Level-2 equations captured the effects of age or vocabulary size on the log-odds of looking to the familiar object over time (same equations as above, with the Subject × Condition random effect and Condition fixed effect removed). Additional models with age, vocabulary size, and the interactions among the three were also computed.

In all models including EVT-2 scores as a measure of vocabulary size were consistently better-fitting models than models with PPVT-4 scores; therefore, EVT-2 scores were used as the sole measure of vocabulary size and models including PPVT-4 will not be discussed further.[2]

---

[2] We also explored whether receptive vocabulary size (PPVT-4 scores) was a significant predictor in the model. EVT-2 and PPVT-4 scores are highly correlated and both are also correlated with age. We included expressive, rather than receptive vocabulary size in the model because EVT-2 scores were a stronger predictor of performance than PPVT-4 scores and provided a better fit of the data for all statistical models calculated. Model fit comparison for CP v. MP trials: simple model with PPVT-GSV: AIC: 1892.7, BIC: 2034.1, full model with PPVT-standard and Age: AIC: 1911.2, BIC: 2120.3. Model fit comparison for NW trials: simple model with PPVT-GSV: AIC: 1195.7, BIC: 1260.1, full model with PPVT-standard and Age: 1186.2 1280.5.

It is computationally and theoretically difficult to estimate the degrees of freedom for mixed-effects model (see Bates, 2006, for discussion of this topic). We therefore analyzed *t*-scores assuming a Gaussian distribution (i.e., $t(\infty)^{\alpha=0.05} > \pm 1.96$ was considered significant).[3]

PREDICTIONS

We expected that, in general, children would look to the familiar object for CP trials and look to the unfamiliar object during NW trials (i.e., disambiguation). We also anticipated that the MP trials would have a significantly different looking pattern compared with the CP trials, indicating that the children were sensitive to the phonological discrepancy between the correct pronunciations and mispronunciations of known words. Furthermore, we expected an interaction between vocabulary size and looking patterns; children with larger vocabulary sizes (as measured by the EVT-2 growth scale values) would look to the familiar object in CP trials and to the unfamiliar image during NW trials more consistently (indexed by the intercept), at a comparatively faster rate (Time) and acceleration (Time^2). Finally, we expected that children with smaller vocabulary sizes would be comparatively less sensitive to the phonological discrepancy of the mispronunciations and would have similar looking patterns for the CP and MP trials.

We predicted a convex-shaped growth curve for CP trials. Although this would be indicated by a more negative parameter estimate, Time^2 was inverted so that a more positive parameter estimate was interpreted as a relatively faster acceleration in the log-odds of looking to the familiar image.

RESULTS

Figure 2 shows the proportion of looks to the familiar object by Condition over time, collapsed across all 34 children. It can be observed that children generally looked to the familiar image upon hearing the correct pronunciation of the object-name and generally looked to the unfamiliar object upon hearing a nonword. For the MP trials, the proportion of looking to the familiar image was consistently close to 0.5, suggesting that the children, as a group, were equally likely to look at either image at any point during the MP trials.

Comparing Looking Patterns during MP Trials v. CP Trials

Two separate mixed effects growth curve models were calculated, each with a single child-level predictor: 1) a model with EVT-2 growth scale values and 2), a model with chronological age. In comparing these models, the model using EVT-2 growth scale values as a predictor was a significantly better fit; both the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were lower for the model with EVT-2 growth scale values as a predictor [AIC:

---

[3] The models were also calculated using an R package that approximated the degrees of freedom using likelihood estimates (Kuznetsova, Brockhoff, & Christensen, 2013), yielding roughly the same *p*-values as listed in the result section.

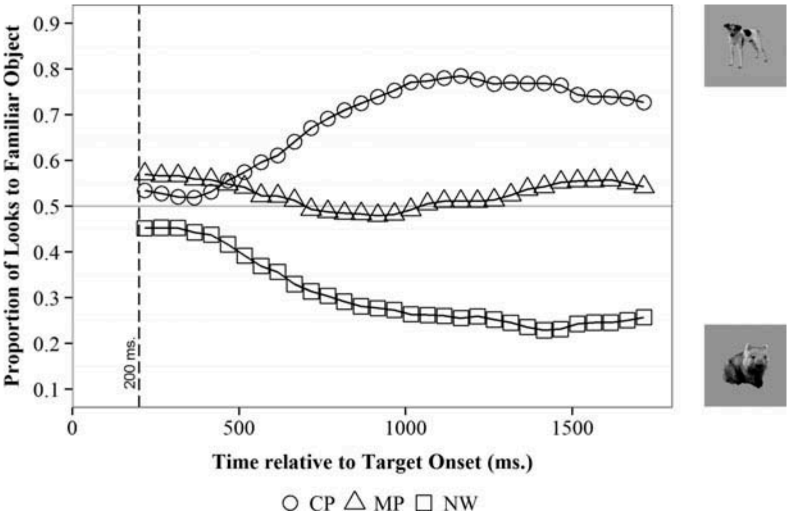


FIGURE 2 Proportion of looks to the familiar object over time, collapsed across children. The zero-point on the abscissa represents the time point at which the onset of the target word occurred. The time frame of the trials used for analysis was 200–1,700 ms after target word onset.

EVT-2 GSV (1,887) < Age (1,892); BIC: EVT-2 GSV (2,028) < Age (2,033). See Appendix for additional information regarding model fit comparison]. Moreover, no parameter estimates involving Age reached significance. Thus, the parameter estimates reported in this section were derived from the model that included EVT-2 growth scale values as a predictor. Finally, a third mixed-effects model including Age, EVT-2 standard scores, and the interaction between the two was also calculated to investigate the findings associated with expressive vocabulary, controlling for chronological age.

The direction of the parameter estimates were the same for all three models: the mixed effects growth curve models confirm that, as a group, the children were sensitive to the mispronounced target word; a main effect of Condition demonstrated that the log-odds of looking to the familiar image was significantly lower for MP trials, relative to CP trials [$\gamma_{0\mathrm{MP}} = -0.98$, $SE = 0.13$, $t = -7.68$, $p < .001$].

Changes in log-odd values as a function of linear time can be thought of as the rate at which the log-odds value changes over time (that is, as differences in slope of the growth curves). As expected, the overall slope for Time was significant [$\gamma_{10} = 2.39$, $SE = 0.36$, $t = 6.66$, $p < .001$], indicating that the log-odds of looking to the familiar image increased over time for CP trials. The interaction between Time and the MP condition was significant; the slope was less steep, relative to the reference (CP) condition [$\gamma_{1\mathrm{MP}} = -2.33$, $SE = 0.48$, $t = -4.83$, $p < .001$]. Indeed, interpreting the parameter estimate relative to the reference (i.e., $\gamma_{10} + \gamma_{1MP}$ or 2.39 – 2.33), the parameter estimate was close to 0, confirming that children were equally likely to look to either the familiar or unfamiliar image over time.

The Time^2 parameter captured the acceleration in change of log-odds. CP trials were positively associated with Time^2 [$\gamma_{20} = 1.24$, $SE = 0.23$, $t = 5.40$, $p < .001$], signifying a faster acceleration in the log-odds of looking to the familiar image over time. MP trials, in comparison, had a more flat curve [$\gamma_{1\mathrm{MP}} = -1.76$, $SE = 0.32$, $t = -5.52$, $p < .001$].

Focusing on the child-level independent variables, there was a significant interaction of Condition and EVT-2 growth scale values; children with larger expressive vocabularies were less likely to look at the familiar object for MP trials, compared to CP trials [$\gamma_{0EVT^{*}MP} = -0.03$, $SE = 0.01$, $t = -2.10$, $p = .04$]. Age as a predictor patterned in the same direction, but was not significant. Regarding the effects of expressive vocabulary on the change of log-odds over time, there was a significant interaction of Time with EVT-2 growth scale values; as the EVT-2 growth scale value increased, the rate of looking to the familiar image for CP trials also increased [$\gamma_{1EVT} = 0.08$, $SE = 0.04$, $t = 2.13$, $p = .03$]. Again, Age patterned in the same direction, but was not significant. Figure 3 depicts the relationship between vocabulary size and looking patterns. As described above, vocabulary size was a continuous variable in the models run, but the children were grouped into three levels of EVT-2 growth scale values in Figures 3 and 4 for the purpose of illustration. It can be observed that children with EVT-2 growth scale values 1 SD above the mean (Figure 3, fourth panel) looked quickly and reliably to the familiar image during CP trials and showed a reliably different looking pattern during MP trials, relative to CP trials. In contrast, the children with EVT-2 growth scale values more than 1 SD below the group mean (Figure 3, second panel) looked less often to the familiar image throughout the duration of the trial and exhibited the least amount of distinction in looking patterns during MP trials, compared to the CP trials.

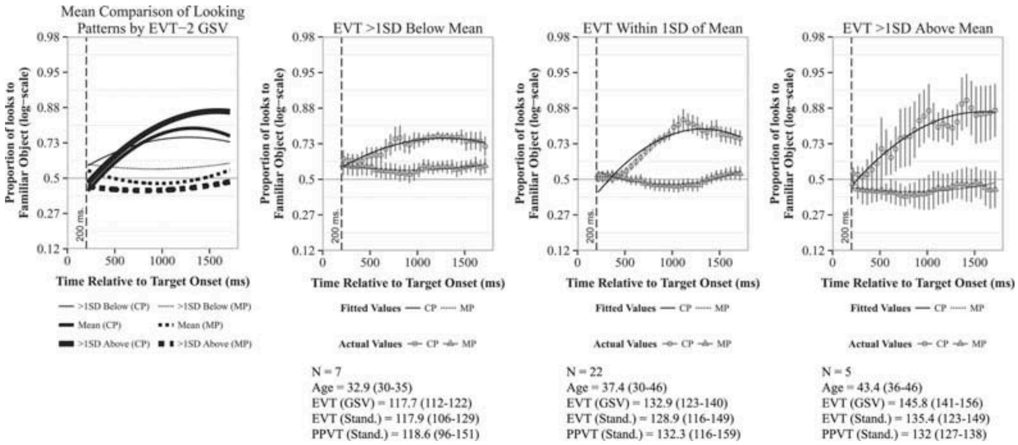


FIGURE 3 Comparison of looking patterns for CP (solid) v. MP (dotted) trials. Participants were grouped into whether their EVT-2 growth scale values (GSV) was greater than 1SD below the group mean, within 1SD of the mean, or greater than 1SD above. Note that the data grouping by EVT-2 GSV is merely for illustrative purposes. The first panel shows differences in model fit, as a function of EVT-2 GSV. The next three panels include model fits, as well as the actual values (i.e., mean proportion and standard error for each time bin) for each group. Children with EVT-2 GSV more than 1SD below the group mean (second panel) showed significantly less differentiation in looking patterns for CP v. MP trials.

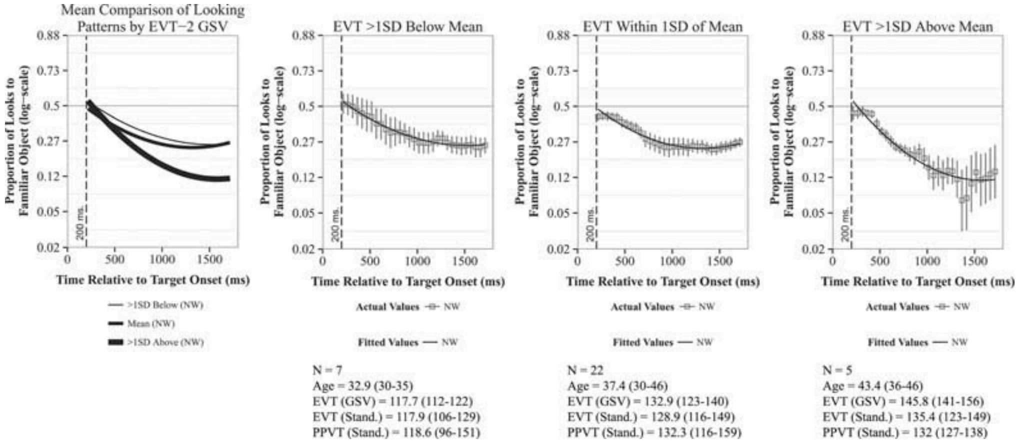


FIGURE 4 Comparison of looking patterns for NW trials. Participants were grouped into whether their EVT-2 growth scale values (GSV) was greater than 1SD below the group mean, within 1SD of the mean, or greater than 1SD above. Note that the data grouping by EVT-2 GSV is merely for illustrative purposes. The first panel shows differences in model fit, as a function of EVT-2 GSV. The next three panels include model fits, as well as the actual values (i.e., mean proportion and standard error for each time bin) for each group. Children with EVT-2 GSV more than 1SD below the group mean (second panel) were less likely to exhibit disambiguation in NW trials.

Although chronological age was not a significant predictor of performance, there was a significant correlation between chronological age and EVT-2 growth scale values [$r = .617$, $p < .001$]. It was of interest to determine whether the results described above attributed to vocabulary size were still present when controlling for age. A third model including Age, EVT-2 standard scores, and the interaction between the two was calculated, and there was a significant main effect of EVT-2 standard scores, but not chronological age; children with larger expressive vocabularies were more likely to look at the familiar image during CP trials, had a faster rate and acceleration in looks to the familiar object for CP trials, and were more sensitive to mispronunciations. This pattern was even more evident for older children with larger vocabularies [$\gamma_{0Age^{*}EVT^{*}MP} = -0.01$, $SE = 0.002$, $t = -2.35$, $p = 0.019$; $\gamma_{2Age^{*}EVT^{*}MP} = -0.02$, $SE = 0.01$, $t = -3.10$, $p = .002$]. This interaction is similar to the findings described in the mixed effects model with only EVT-2 growth scale values, in that the older children with higher EVT-2 standard scores were also the children with the higher growth scale values. This more complex model was not a better fit, compared to the model with only EVT-2 GSV [AIC: EVT-2 GSV (1,887) < EVT-2 standard score (1,889); BIC: EVT-2 GSV (2,028) < EVT-2 standard score (2,098)]. This suggests that, despite the correlation between chronological age and EVT-2 growth scale values, a model including only EVT-growth scale values was sufficient to estimate the eye gaze patterns of the CP and MP trials.

Looking Patterns during NW Trials

The NW trials were analyzed separately to examine the looking patterns when a nonword was heard. Again, two separate mixed-effects models were compared in order to explore the effects of the child-level independent variables on the looking patterns. A comparison of the two models again confirmed that EVT-2 growth scale values, rather than chronological age, provided a better fit of eye gaze patterns for NW trials; the parameter estimates including age were not significant, and the AIC and BIC were slightly lower for the EVT-2 growth scale value model [AIC: EVT-2 GSV (1,187) < Age (1,188); BIC: EVT-2 GSV (1,251) < Age (1,253). See Appendix for additional information regarding model fit comparison)]. A model including both chronological age and EVT-2 standard scores was also calculated to investigate the relationship between expressive vocabulary and looking patterns during NW trials, while controlling for chronological age.

The mixed-effects models with EVT-2 growth scale values revealed that children were more likely to look at the unfamiliar objects (indicated by negative parameter estimates) during NW trials over time [Intercept: $\gamma_{00} = -0.93$, $SE = 0.1$, $t = -9.52$, $p < .001$, Time: $\gamma_{10} = -1.98$, $SE = 0.38$, $t = -5.21$, $p < .001$; Time^2: $\gamma_{20} = -0.9$, $SE = 0.27$, $t = -3.31$, $p < .001$]. Regarding the child-level variables, children with larger expressive vocabularies were even more likely to look at the unfamiliar object upon hearing a nonsense word [$\gamma_{0EVT} = -0.03$, $SE = 0.01$, $t = -3.28$ $p < .001$]. Figure 4 illustrates the relationship between EVT-2 growth scale values and looking patterns to the unfamiliar image.

The results of a third model with age, EVT-2 standard scores, and the interaction between the two also confirmed that EVT-2 standard scores were associated with more overall looks to the unfamiliar image [$\gamma_{0EVT} = -0.03$, $SE = 0.01$, $t = -4.03$, $p < .001$], as well as a faster rate and acceleration in looking to the unfamiliar image [Time: $\gamma_{1EVT} = -0.06$, $SE = 0.03$, $t = -2.29$, $p < .022$; Time^2: $\gamma_{2EVT} = -0.07$, $SE = 0.02$, $t = -3.34$, $p < .001$]. The parameter estimate for age did not reach significance, but there was a significant three-way interaction; older children with larger expressive vocabularies had a faster rate of looking to the unfamiliar image relative to younger children with high vocabularies [$\gamma_{1Age*EVT} = -0.03$, $SE = 0.01$, $t = -4.47$, $p < .001$]. In comparing the AIC and BIC of the model including only EVT-2 growth scale values and the more complex model including both age and EVT-2 standard scores, it is unclear which model provides the best fit of the data [AIC: EVT-2 GSV (1,187) > EVT-2 standard scores (1,162); BIC: EVT-2 GSV (1,251) < EVT-2 standard scores(1,257)]. However the more complex model suggests a significant association with vocabulary size and the rate of acceleration of looks to the unfamiliar image when controlling for chronological age. This was not apparent in the model including only EVT-2 growth scale values.

DISCUSSION

This study was designed to explore the relationship between vocabulary size and lexical processing efficiency in children from 30 to 46 months of age. There were three main results of this study. First, we found that expressive vocabulary size was associated with eye gaze patterns for both familiar words and mispronunciations of familiar words. Children with larger vocabularies were faster at identifying familiar words and were more sensitive to incorrect productions of familiar words. That is, they more quickly and consistently looked to novel objects rather than to familiar objects when they heard a one-feature mispronunciation of a familiar word. We also found that expressive vocabulary size was associated with eye gaze patterns to unfamiliar objects for nonword stimuli. Children with larger vocabularies looked more quickly and consistently to the picture of an unfamiliar object when they heard an unfamiliar word relative to children with smaller vocabularies.

The results of this study differ from the results of some previous research studies using the LWL paradigm in several respects. Unlike previous studies (Fernald et al., 2006; Marchman & Fernald, 2008), we did not observe a main effect of age on eye gaze patterns for any of the three conditions of correct productions, mispronunciations, or nonwords. The mixed-effects models including both age and EVT-2 standard scores lend further support to this finding. First, the main effect of EVT-2 standard scores remained significant, even when age was included in the model. Second, the group of children with EVT-2 standard scores greater than 1 SD above the mean included a large age range (N = 7, mean EVT-2 (stand.): 144.7 mean age in months: 34.7, range: 30-45). Similarly, the group of children with EVT-2 standard scores more than 1 SD below the mean also included a large age range (N = 7, mean EVT-2 (stand.): 113.4; mean age in months: 37.4, range: 34-42). The significant interaction between age and EVT-2 standard scores bolster the results observed regarding the relationship between vocabulary size and lexical processing. Given any particular EVT-2 standard score (e.g., EVT-2 standard score = 100), older children with this score would have a higher GSV (e.g., larger vocabularies) than younger children with this same score. As was demonstrated by the model that included EVT-2 GSV scores, children with larger vocabularies exhibited better performance than children with smaller vocabularies.

There are a number of reasons why the results of this study differ from previous research. Specifically, we used a different experimental paradigm (pictures of one familiar and one unfamiliar objects rather than pictures of two familiar objects), we relied on a different way of estimating expressive vocabulary size (EVT-2, a direct measure of expressive vocabulary size, rather than the MCDI, a parent report), and we used a different method of analysis (growth curve analysis rather than a latency/accuracy comparison). Furthermore, this study focused on older children (30- to 46-month-olds rather than 15- to 25-month-olds) and included a larger age range than most previous studies. Furthermore, the age range was continuous from 30 to 46 months, rather than grouping children into one-month age groups separated by three to seven months, as has been done in previous research. Having a relatively large range of age and vocabulary sizes made it possible to examine whether vocabulary size or age was a better predictor of lexical processing efficiency in this study. The fact that expressive vocabulary, but not age, was a significant predictor of looking patterns for all three conditions in this study suggests that the effects of age for younger children that were observed in previous studies may be related to non-linguistic age-related factors, such as attention.

The finding that children with larger vocabularies were more sensitive to one-feature mispronunciations than peers with smaller vocabularies provides additional experimental support for the claim that increases in vocabulary size result in more segmental, phonological representations in the lexicon (Curtin & Werker, 2007; Edwards, Beckman, & Munson, 2004; Werker & Curtin, 2005). Whereas Mayor and Plunkett (2014) found a similar effect of vocabulary size on the sensitivity to mispronunciation in a computational model based on the TRACE model of speech perception (McClelland & Elman, 1986), previous studies using this paradigm either did not find an effect of vocabulary size on eye gaze patterns for mispronunciations (Bailey &

Plunkett, 2002; Ballem & Plunkett, 2005; Mani & Plunkett, 2007; Swingley & Aslin, 2000, 2002) or did not expressly examine this relationship (White & Morgan, 2008).

There are at least three possible reasons why a relationship between vocabulary size and eye gaze patterns for mispronunciations was observed in this study but not in previous studies. First, the participants in this study were older: our participants were 30 to 46 months of age, whereas participants in the previous studies ranged from 14 to 24 months of age. It is possible that a larger range of vocabulary size than was included in previous studies is necessary to observe the effect of vocabulary size on the response to mispronunciations. Second, the current study used growth curve analysis to analyze the results, rather than examining latency and accuracy; growth curve analysis is a more powerful statistical technique for examining such relationships (Mirman et al., 2008; Mirman, 2014). A replication or reanalysis of data from the earlier studies would be needed to confirm or reject this explanation. Finally, the studies that did not find an effect of vocabulary size presented pictures of two familiar objects (Bailey & Plunkett, 2002; Ballem & Plunkett, 2005; Mani & Plunkett, 2007; Swingley & Aslin, 2000, 2002), whereas the current study, as in White and Morgan (2008), presented pictures of a familiar and an unfamiliar object. As noted above, if both pictures are of familiar objects, then a listener will be biased toward looking to the image representing the target word when the listener hears a mispronunciation. By contrast, that bias is not present in the current paradigm; when a listener hears a mispronunciation, the listener must decide whether it is a production of the target word or a novel word, perhaps allowing this paradigm to be more sensitive to detecting a relationship between vocabulary size and responses to mispronunciations. However, it should be noted that there is a limitation to both the White and Morgan (2008) paradigm and the paradigm used in the current study. In both studies, different sets of image pairs were used for the correct pronunciation/mispronunciation conditions and the nonword condition. Nonwords were never presented in the correct pronunciation/mispronunciation conditions and names of familiar objects were never presented in the nonword condition. This design meant that it was necessary to analyze the nonword condition separately from the correct pronunciation/mispronunciation conditions. Therefore, it was not possible to compare the correct pronunciation and nonword conditions directly. Based on eye gaze patterns at baseline, there did not appear to be a looking preference toward either the novel or familiar image, but it would be desirable to be able to evaluate this statistically during the time period after target word onset. However, that was not possible in this study, given the design used.

The relationship between vocabulary size and eye gaze patterns in the novel word condition is similar to the results of Bion et al. (2013) for 24- and 30-month-olds. Although infants have a bias to associate a novel object and a novel name well before they know many words, the results of this study, in combination with those of Bion et al. suggest that by age 2, N3C/mutual exclusivity has become linked to vocabulary size; children with larger vocabularies more efficiently associate novel objects with novel names.

The current study provides support for the growing body of evidence that differences in vocabulary size in young children are associated with differences in online processing of both familiar and novel words. Critically, these findings illustrate the interdependence of vocabulary size and vocabulary growth (Elman et al., 1996, pp. 124–129; Plunkett, Sinha, Møller, & Strandsby, 1992). Given that the average speaking rate is about two syllables per second, even small differences in lexical processing speed place some children at an advantage and others at a disadvantage with respect to word learning in particular and language learning more generally. Children with larger vocabularies process familiar words more efficiently, thus leaving them with more time and greater cognitive resources for other aspects of language comprehension. In addition, children with larger vocabularies are also more efficient at differentiating between familiar words and novel words that differ by only a single phonological feature, and they are better at disambiguating novel words. Children who can quickly and reliably recognize a novel word as a "new" word will more rapidly add new words to their vocabulary. Unfortunately for children with small vocabularies, the results of this study are just one more example of the "Matthew Effect" in child language development. That is, children with larger vocabularies process both familiar and novel words more efficiently than children with smaller vocabularies, thus helping to maintain or even increase differences in vocabulary size over time.

ACKNOWLEDGMENTS

We thank Mary E. Beckman, Jianshen Chen, David Kaplan, Danielle Lee, Tristan Mahr, Morgan Meredith, Benjamin Munson, Pat Reidy, Erica Richmond, Alissa Schneeberg, Nancy Wermuth, and other members of the Learning To Talk Laboratory for help with many aspects of this study. We also thank the children who participated and their parents.

FUNDING

This research was supported by NIDCD Grant R01-02932 to Jan Edwards, Mary E. Beckman, and Benjamin Munson, NICHD Grant 2-T32-HD049899 to Maryellen MacDonald, and NICHD Grant P30-HD03352 to the Waisman Center.

REFERENCES

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*, 419–439.

Bailey, T. M., & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development*, *17*, 1265–1282. doi:10.1016/S0885-2014(02)00116-8

Ballem, K. D., & Plunkett, K. (2005). Phonological specificity in children at 1;2. *Journal of Child Language*, *32*, 159–173. doi:10.1017/S0305000904006567

Barr, D. J. (2008). Analyzing "visual world" eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *59*, 457–474. doi:10.1016/j.jml.2007.09.002

Bates, D. (2006). Lmer, p-values and all that. *ETH Seminar for Statistics*. Retrieved from https://stat.ethz.ch/pipermail/r-help/2006-May/094765.html

Bates, D., Maechler, M., & Bolker, B. (2013). lme4: Linear mixed-effects models using S4 classes. Retrieved from http://cran.rproject.org/package=lme4

Beckage, N. M., Smith, L. B., & Hills, T. (2010). Semantic network connectivity is related to vocabulary growth rate in children trajectories of early vocabulary growth. In *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society 2010* (pp. 2769–2774).

Benedict, H. (2008). Early lexical development: Comprehension and production. *Journal of Child Language*, *6*, 183–200. doi:10.1017/S0305000900002245

Bion, R. H., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word-object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, *126*, 39–53. doi:10.1016/j.cognition.2012.08.008

Charles-Luce, J., & Luce, P. A. (2009). Similarity neighbourhoods of words in young children's lexicons. *Journal of Child Language*, *17*, 205–215. doi:10.1017/S0305000900013180

Colunga, E., & Sims, C. E. (2011). Early talkers and late talkers know nouns that license different word learning biases. In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society 2011* (pp. 2550–2555).

Colunga, E., & Sims, C. E. (2012). Early-talker and late-talker toddlers and networks show different word learning biases. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society 2012* (pp. 246–251).

Curtin, S., & Werker, J. F. (2007). The perceptual foundations of phonological development. In M. G. Gaskell, G. T. M. Altmann, P. Bloom, A. Caramazza, & P. Levelt (Eds.), *Oxford handbook of psycholinguistics* (pp. 579–599). Oxford, England: Oxford University Press.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, *28*, 125–127. doi:10.3758/BF03203646

Dewar, K., & Xu, F. (2007). Do 9-month-old infants expect distinct words to refer to kinds? *Developmental Psychology*, *43*, 1227–1238. doi:10.1037/0012-1649.43.5.1227

Diesendruck, G., Gelman, S. A., & Lebowitz, K. (1998). Conceptual and linguistic biases in children's word learning. *Developmental Psychology*, *34*, 823–839. doi:10.1037//0012-1649.34.5.823

Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology*, *37*, 630–641. doi:10.1037//0012-1649.37.5.630

Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test* (4th ed.). Minneapolis, MN: NCS Pearson.

Duta, M. D., Styles, S. J., & Plunkett, K. (2012). ERP correlates of unexpected word forms in a picture-word study of infants and adults. *Developmental Cognitive Neuroscience*, *2*, 223–234. doi:10.1016/j.dcn.2012.01.003

Edwards, J., Beckman, M. E., & Munson, B. (2004). Vocabulary size and phonotactic production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research*, *47*, 421–436. doi:10.1044/1092-4388(2004/034)

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D. J., Bates, E., Hartung, J. P., . . . Reilly, J. S. (1993). *The MacArthur communicative development inventories: User's guide and technical manual*. Baltimore, MD: Paul H. Brokes.

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18. *Developmental Science*, *16*, 234–248. doi:10.1111/desc.12019

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

Fernald, A., Swingley, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development*, *72*, 1003–1015. doi:10.1111/1467-8624.00331

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. M. Fernández, & H. Clahsen (Eds.), *Developmental psycholonguistics: On-line methods in children's language processing*. Amsterdam, the Netherlands: John Benjamins.

Friedrich, M., & Friederici, A. D. (2005). Phonotactic knowledge and lexical-semantic processing in one-year-olds: brain responses to words and nonsense words in picture contexts. *Journal of Cognitive Neuroscience*, *17*, 1785–1802. doi:10.1162/089892905774589172

Goldfield, B. A., & Reznick, J. S. (2009). Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of Child Language*, *17*, 171. doi:10.1017/S0305000900013167

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology*, *28*, 99–108. doi:10.1037//0012-1649.28.1.99

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

Graf Estes, K., Edwards, J., & Saffran, J. R. (2011). Phonotactic constraints on infant word learning. *Infancy*, *16*, 180–197. doi:10.1111/j.1532-7078.2010.00046.x

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*, *28*, 267–283. doi:10.3758/BF03204386

Haith, M. M., Wentworth, N., & Canfield, R. (1993). The formation of expectations in early infancy. *Advances in Infancy Research*, *8*, 251–297.

Hall, D. G. (1991). Acquiring proper nouns for familiar and unfamiliar animate objects: Two-year-olds' word-learning biases. *Child Development*, *62*, 1142–1154. doi:10.1111/j.1467-8624.1991.tb01595.x

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H. Brookes.

Heibeck, T. H., & Markman, E. M. (1987). Word learning in children: An examination of fast mapping. *Child Development*, *58*, 1021–1034. doi:10.2307/1130543

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*, 1368–1378. doi:10.1111/1467-8624.00612

Hurtado, N., Grüter, T., Marchman, V. A., & Fernald, A. (2014). Relative language exposure, processing efficiency and vocabulary in Spanish–English bilingual toddlers. *Bilingualism: Language and Cognition*, *17*, 189–202. doi:10.1017/S136672891300014X

Inhoff, A. W., & Radach, R. (1998). Definition and computtion of oculomotor measures in the study of cognitive processes. In G. M. Underwood (Ed.), *Eye guidance in reading and scene peception* (pp. 29–53). Oxford, England: Elsevier.

Jones, S. S., & Smith, L. B. (2005). Object name learning and object perception: A deficit in late talkers. *Journal of Child Language*, *32*, 223–240. doi:10.1017/S0305000904006646

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2013). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). Retrieved from http://cran.r-project.org/package=lmerTest

Luce, P. A., Goldinger, S. D., Auer, Jr., E. T., & Vitevitch, M. S. (2000). Phonetic priming, neighborhood activation, and PARSYN. *Perception & Psychophysics*, *62*, 615–625. doi:10.3758/BF03212113

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*, 1–36.

Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, *57*, 252–272. doi:10.1016/j.jml.2007.03.005

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*, F9–16. doi:10.1111/j.1467-7687.2008.00671.x

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, *14*, 57–77. doi:10.1016/0364-0213(90)90026-S

Markman, E. M. (1991). The whole object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. In J. P. Byrnes & S. A. Gelman (Eds.), *Perspectives on language and cognition: Interrelations in development* (pp. 72–106). Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511983689.004

Markman, E. M., Wasow, J. L., & Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology*, *47*, 241–275. doi:10.1016/S0010-0285(03)00034-3

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word recognition. *Cognition*, *25*, 71–102. doi:10.1016/0010-0277(87)90005-9

Marslen-Wilson, W. D., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review*, *101*, 653–675. doi:10.1037/0033-295X.101.4.653

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*, 29–63. doi:10.1016/0010-0285(78)90018-X

Mather, E., & Plunkett, K. (2009). Learning words over time: The role of stimulus repetition in mutual exclusivity. *Infancy*, *14*, 60–76. doi:10.1080/15250000802569702

Mather, E., & Plunkett, K. (2010). Novel labels support 10-month-olds' attention to novel objects. *Journal of Experimental Child Psychology*, *105*, 232–242. doi:10.1016/j.jecp.2009.11.004

Mather, E., & Plunkett, K. (2011). Mutual exclusivity and phonological novelty constrain word learning at 16 months. *Journal of Child Language*, *38*, 933–950. doi:10.1017/S0305000910000401

Mayor, J., & Plunkett, K. (2014). Infant word recognition: Insights from TRACE simulations. *Journal of Memory and Language*, *71*, 89–123. doi:10.1016/j.jml.2013.09.009

McCall, R. B., & Carriger, M. S. (1993). A meta-analysis of infant habituation and recognition memory performance as predictors of later IQ. *Child Development*, *64*, 57–79. doi:10.1111/1467-8624.ep9309015107

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86. doi:10.1016/0010-0285(86)90015-0

McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science*, *317*, 631. doi:10.1126/science.1144073

Merriman, W. E., Bowman, L. L., & MacWhinney, B. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development*, *54*, 1–132. doi:10.2307/1166130

Merriman, W. E., & Schuster, J. M. (1991). Young children's disambiguation of object name reference. *Child Development*, *62*, 1288–1301. doi:10.1111/j.1467-8624.1991.tb01606.x

Mervis, C. B., & Bertrand, J. (1994). Acquisition of the novel name–nameless category (N3C) principle. *Child Development*, *65*, 1646–1662. doi:10.1111/j.1467-8624.1994.tb00840.x

Mirman, D. (2014). *Growth curve analysis and visualization using R*. Boca Raton, FL: Chapman & Hall/CRC Press, Taylor & Francis Group.

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, *59*, 475–494. doi:10.1016/j.jml.2007.11.006

Mirman, D., Yee, E., Blumstein, S. E., & Magnuson, J. S. (2011). Theories of spoken word recognition deficits in aphasia: Evidence from eye-tracking and computational modeling. *Brain and Language*, *117*, 53–68. doi:10.1016/j.bandl.2011.01.004

Moore, C., Angelopoulos, M., & Bennett, P. (1999). Word learning in the context of referential and salience cues. *Developmental Psychology*, *35*, 60–68. doi:10.1037/0012-1649.35.1.60

Moulines, E., & Charpentier, F. (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication*, *9*, 453–467. doi:10.1016/0167-6393(90)90021-Z

Paul, R. (1993). Patterns of development in late talkers: Preschool years. *Communication Disorders Quarterly*, *15*, 7–14. doi:10.1177/152574019301500103

Plunkett, K., Sinha, C., Møller, M. F., & Strandsby, O. (1992). Symbol grounding or the emergence of symbols? Vocabulary growth in children and a connectionist net. *Connection Science*, *3-4*, 293–312. doi:10.1080/09540099208946620

Psychology Software Tools, Inc. (2010). E-Prime® 2.0 Professional (Build 2.0.8.90) [Computer Software]. Pittsburgh, PA.

Radach, R., Heller, D., & Inhoff, A. (1999). Occurrence and function of very short fixation durations in reading. In W. Becker, H. Deubel, & T. Mergner (Eds.), *Current oculomotor research: Physiological and psychological aspects* (pp. 321–331). New York, NY: Springer. doi:10.1007/978-1-4757-3054-8_46

Rescorla, L., Mirak, J., & Singh, L. (2000). Vocabulary growth in late talkers: Lexical development from 2;0 to 3;0. *Journal of Child Language*, *27*, 293–311. doi:10.1017/S030500090000413X

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools.

Sheng, L., & McGregor, K. K. (2010). Lexical-semantic organization in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *53*, 146–159. doi:10.1044/1092-4388(2009/08-0160)

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

Smith, L. B., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*, 1558–1568. doi:10.1016/j.cognition.2007.06.010

Smits, R., Warner, N., McQueen, J. M., & Cutler, A. (2003). Unfolding of phonetic information over time: a database of Dutch diphone perception. *The Journal of the Acoustical Society of America*, *113*, 563–574. doi:10.1121/1.1525287

Storkel, H. L. (2001). Learning new words: phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research*, *44*, 1321–1337. doi:10.1044/1092-4388(2001/103)

Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech: A working model. *Journal of Phonetics*, *39*, 456–466. doi:10.1016/j.wocn.2010.09.001

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*, 147–166. doi:10.1016/S0010-0277(00)00081-0

Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, *13*, 480–484. doi:10.1111/1467-9280.00485

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing of spoken words. *Psychological Science*, *9*, 325–329. doi:10.1111/1467-9280.00064

Vouloumanos, A. (2008). Fine-grained sensitivity to statistical information in adult word learning. *Cognition*, *107*, 729–742. doi:10.1016/j.cognition.2007.08.007

Walley, A. C. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review*, *13*, 286–350. doi:10.1006/drev.1993.1015

Wass, S. V., Smith, T. J., & Johnson, M. H. (2013). Parsing eye-tracking data of variable quality to provide accurate fixation duration estimates in infants and adults. *Behavior Research Methods*, *45*, 229–250. doi:10.3758/s13428-012-0245-6

Waxman, S. R., & Booth, A. (2003). The origins and evolution of links between word learning and conceptual organization: New evidence from 11-month-olds. *Developmental Science*, *6*, 128–135. doi:10.1111/1467-7687.00262

Waxman, S. R., & Lidz, J. L. (2006). Early word learning. In D. Kuhn & R. Siegler (Eds.), *Handbook of Child Psychology* (6th ed., pp. 299–335). Hoboken NJ: Wiley.

Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development*, *1*, 197–234. doi:10.1080/15475441.2005.9684216

Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy*, *3*, 239–273. doi:10.1207/15250000252828226

White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, *59*, 114–132. doi:10.1016/j.jml.2008.03.001

Williams, K. T. (2007). *Expressive vocabulary test* (2nd ed.). Minneapolis, MN: Pearson Assessments.

APPENDIX

Measures of Model Fit, Deviance, and Variance of Random Effects of All Mixed-Effects Models

	CP v. MP						***NW***					
	Age		***EVT-GSV***		***Age * EVT***		***Age***		***EVT-GSV***		***Age * EVT***	
Random Effects	**Variance**	**SD**	**Variance**	**SD**	**Variance**	**SD**	**Variance**	**SD**	**Variance**	**SD**	**Variance**	**SD**
Child												
U_{0j} (Intercept)	0.096	0.311	0.104	0.323	0.101	0.318	0.419	0.647	0.321	0.567	0.282	0.531
U_{1j} (Time)	0.517	0.719	0.426	0.653	0.423	0.650	4.900	2.214	4.763	2.182	2.907	1.705
U_{2j} (Time^2)	0.037	0.192	0.046	0.214	0.051	0.225	2.366	1.538	2.407	1.551	1.564	1.251
Child* Condition												
$W_{0(j,k)}$ (Intercept)	0.287	0.536	0.272	0.522	0.229	0.478	–	–	–	–	–	–
$W_{1(j,k)}$ (Time)	4.062	2.015	3.860	1.965	3.641	1.908	–	–	–	–	–	–
$W_{2(j,k)}$ (Time^2)	1.717	1.311	1.638	1.280	1.383	1.176	–	–	–	–	–	–
$R_{t(j,k)}$	0.098	0.313	0.098	0.313	0.098	0.313	0.126	0.354	0.126	0.354	0.126	0.354
df (Fixed Effects)	25		25		37		13		13		19	
AIC	1892		1887		1889		1188		1187		1162	
BIC	2033		2028		2098		1253		1251		1257	
Deviance	1842		1837		1815		1162		1161		1124	
Deviance (REML)	1879		1882		1949		1180		1183		1193	
Log-Likelihood	−921		−918.3		−907.5		−581.1		−580.3		−562.1	