Developmental Psychology
2007, Vol. 43, No. 2, 454–464

0012-1649/07/$12.00 DOI: 10.1037/0012-1649.43.2.454

Lexical Exposure and Word-Form Encoding in 1.5-Year-Olds

Daniel Swingley
University of Pennsylvania

In this study, 1.5-year-olds were taught a novel word. Some children were familiarized with the word's phonological form before learning the word's meaning. Fidelity of phonological encoding was tested in a picture-fixation task using correctly pronounced and mispronounced stimuli. Only children with additional exposure in familiarization showed reduced recognition performance given slight mispronunciations relative to correct pronunciations; children with fewer exposures did not. Mathematical modeling of vocabulary exposure indicated that children may hear thousands of words frequently enough for accurate encoding. The results provide evidence compatible with partial failure of phonological encoding at 19 months of age, demonstrate that this limitation in learning does not always hinder word recognition, and show the value of infants' word-form encoding in early lexical development.

Keywords: lexical development, word learning, word recognition, phonology

Even before infants know what any words mean, their parents speak to them. These interactions are often singsong parental monologues for which the communicative content, from the perspective of the infant, is carried by intonation and tone of voice (e.g., Ferguson, 1964; Fernald, 1989; Fernald & Simon, 1984). Although these monologues convey a limited range of meanings, they are linguistically important in providing the foundation for the child's phonology and lexicon. Both phonological and lexical development begin in infancy, as children fine-tune their initial speech perception capacities in the service of recognizing familiar words (e.g., Aslin, Jusczyk, & Pisoni, 1998; Kuhl, 2004; Werker & Yeung, 2005).

Early learning of words' forms may play a direct role in the ontogenesis of the vocabulary. Children are usually considered to know a word only when they seem aware of its meaning, so vocabulary development is generally viewed as a matter of accumulating sound–meaning correspondences. In keeping with this perspective, experimental studies of word learning usually introduce a novel word form and a novel concept at the same time, so that success demands encoding of the form, the category, and the link between them (Woodward & Markman, 1998). Yet for the words of the early vocabulary, this situation is probably not the modal case. By 8 months, infants may already be familiar with several hundred word forms (Swingley, 2005b). For these words, vocabulary building is a matter of linking existing phonological forms with their denotations (which may or may not be familiar categories).

These considerations motivated the present research, which examined whether word learning is easier when children are already familiar with the word's form. One possibility is that, in the 2nd year, when children begin linking many phonological forms with meanings, phonological encoding no longer presents much of a barrier to learning, and children's word learning is limited primarily by their ability to surmise the meanings of novel words. Alternatively, children may find words easier to learn when the forms of those words are already known. In the present study, this question was addressed by familiarizing children with the form of a word only, teaching children what the word referred to, and then testing their learning relative to children who had not been previously familiarized with the taught word.

Another question addressed concerned children's ability to accurately learn the phonological forms of words, permitting differentiation of similar but distinct words.[1] One conclusion of recent work on phonological encoding in speech perception is that the words most children tend to learn early, such as *shoe*, *baby*, and *Mommy*, are stored with their features intact but that novel words taught in laboratory settings often appear to be missing some phonological information. Most of the research supporting full phonological encoding for well-known words has used a mispronunciation method, revealing that infants and children recognize words less well when those words are mispronounced (for infants, Bortfeld, Morgan, Golinkoff, & Rathbun, 2005; Swingley, 2005a; Vihman, Nakai, DePaolis, & Hallé, 2004; see also Hallé & de Boysson-Bardies, 1996; Jusczyk & Aslin, 1995; for 14- to 24-month-olds, Bailey & Plunkett, 2002; Ballem & Plunkett, 2005;

This work was supported by grants from the Max-Planck Society, the Dutch National Science Foundation (NWO), and U.S. National Science Foundation Grant 0433567 to Delphine Dahan and Daniel Swingley. Portions of the research reported here were presented at the Boston University Conference on Language Development in 2002. I thank Mark Liberman for discussion of the species trapping problem and other matters bearing on word frequency estimation. I also thank the staff at the Nijmegen babylab for their unstinting efforts (in particular, Angela Khadar and Margret van Beuningen) and the many parents and children who made this work possible through their participation.

Correspondence concerning this article should be addressed to Daniel Swingley, Department of Psychology, University of Pennsylvania, 3401 Walnut Street 302C, Philadelphia, PA 19104. E-mail: swingley@psych.upenn.edu

[1] The word *phonological* is used here to refer to any knowledge of the sound form of words. This is not intended to be read as a claim that children's words are made up of contrastive phonological features or that children necessarily ascribe contrastiveness to all phonologically different strings.

Fennell & Werker, 2003, 2004; Nazzi, 2005; Swingley, 2003; Swingley & Aslin, 2000, 2002; White, Morgan, & Wier, 2004). Mispronunciation effects indicate that children know the phonological form of a word well enough to treat the canonical form as a more typical realization than the deviant form. In several of these studies, reduced recognition performance for mispronounced words has been shown to be unrelated to children's vocabulary sizes or their ability to say the tested words correctly, suggesting a general capacity for successful encoding—at least when children are tested on words like *Mommy* and *dog*, words they have likely known as long as they have known any words at all.

However, tests of children's phonological encoding in the learning of new words have revealed persistent difficulties, particularly among children younger than 15 months old. Early studies showing children's failure to learn similar-sounding pairs of words may have underestimated performance by using relatively demanding object-choice tasks (e.g., Garnica, 1973), but even when task demands are minimized, 14-month-olds sometimes fail to demonstrate full phonological encoding of novel words. Several studies supporting this conclusion have been conducted by Werker and her colleagues using an audiovisual habituation procedure (e.g., Stager & Werker, 1997; Werker, Fennell, Corcoran, & Stager, 2002). In this task, children are habituated to one or two word–object pairings and tested for their detection of violations of these pairings. Detection, as reflected in increased looking times, provides evidence for children's memory for the trained associations between word and object. Fourteen-month-olds typically succeed when the words are dissimilar and fail when the words differ by one or two phonological features (e.g., *bin*, *din*; Pater, Stager, & Werker, 2004; though see Fennell, 2006).

Children participating in the audiovisual habituation procedure often succeed by about 17 months of age, suggesting that the fragile phonological encoding or analysis shown in 14-month-olds abates to some degree by 1.5 years of age. Nevertheless, the apparent discrepancy between performance on newly learned and long-familiar words in children of about 14 months raises the question of the generality of the findings for both new and familiar words. In addition, the 17-month-olds' success in the habituation task followed 40–120 repetitions of the words, all immediately before the test.[2] Determining whether this many exposures is necessary for learning will help guide intuitions about the ease with which children can encode words successfully: Do young children store most words with full phonological descriptions (permitting differentiation of phonological variants) or only the few words with which they are most familiar?

One recent study (Ballem & Plunkett, 2005) examined children's sensitivity to mispronunciation of newly taught words using a picture-fixation task. Fourteen-month-olds were shown teaching displays in which two novel words were used to name two novel objects. At test, children recognized these novel words upon hearing correct pronunciations (in the second of two trial blocks) but not upon hearing mispronunciations (in either trial block). However, children's responses to the correct pronunciations and mispronunciations were not significantly different, making conclusions difficult to draw with certainty.

When children respond in these tasks as though they have not perceived, encoded, or otherwise interpreted a phonological distinction the way older children or adults do, how one accounts for this behavior depends in part on one's view of phonological knowledge, because different views of phonological knowledge suggest different descriptions of imperfect learning. Some accounts of phonological representation posit that words are represented by a set of features that take on discrete (usually binary) values. For example, a consonant may be specified in part by whether it is [+voiced] or [−voiced]. If these phonological features are taken to embody the child's knowledge of how a word sounds, then incomplete encoding is a failure to determine the values of one or more features.

A related possibility is that words are stored as probabilities of phonological feature values and that incomplete encoding is manifested as flat (or not sharply peaked) probability distributions over features. This account demands more complex phonological representations but allows for gradual learning of lexical forms rather than leaps from unspecified to fully specified features.

Other theories of lexical representation suggest alternative conceptualizations of developmental change in phonological encoding and interpretation. For example, 1- and 2-year-olds may not encode words as hierarchies of discrete phonological features (or even probabilities over features) because such features themselves are emergent consequences of later lexical development (Metsala & Walley, 1998; Storkel, 2002). Or children might not encode words as sets of discrete features because, in fact, even in adults, words are stored as atomic, unanalyzed phonetic objects, as proposed by some exemplar theories of lexical representation (e.g., Johnson, 1997). If words are stored without reference to sublexical components, increasing specification of a word's form is a matter of improving one's estimate of the most typical realization of the word and of the word's range of phonetic variation. The present study was not intended to differentiate these ways of thinking about how sounds are represented in words; rather, the study examined effects of children's familiarity with a particular phonological form on the learning and representation of that form when it is linked to a specific denotation. The results are discussed from the perspective of a range of generic types of phonological representation.

Children 1.5 years old were taught a novel word under two exposure conditions: one in which children heard the word several times in a semantically neutral context before being taught its meaning, and one in which children were not given this preexposure (and were instead preexposed to a foil word). All children were then tested for their recognition of the taught novel word in a picture-fixation procedure. On some trials, children heard the word pronounced correctly, as it had been taught; on other trials, children heard the word mispronounced, to probe for the decrements in recognition that indicate successful encoding.

Method

Children from 18 to 20 months of age were exposed to 14 repetitions of a novel bisyllabic word in a story. The story provided little or no constraint on the meaning of the word beyond its being a noun. After the story phase of the experiment, children in the preexposure condition were shown a novel object and told, in eight sentences, that the word that had been in the story was the name of this novel object. The other children, in the control condition, were

[2] This range of repetitions reflects the minimum and maximum given the experimental parameter settings described in Werker et al. (2002).

instead taught a new novel word, which did not match the novel word of the story they had heard. Thus, all children were taught a novel word, but only half of the children had been prefamiliarized with the word's form.

The test phase followed this teaching phase. Children's recognition of the novel word was tested in a series of trials on which children were shown the novel object and a distracter object side by side and asked about the novel object by name. On some of these trials, the novel word was mispronounced. Children's eye movements were monitored as the words were spoken, and fixation to the named picture was used as the dependent measure of word recognition. Children's performance on correct-pronunciation (CP) trials indexed their learning of the link between the sound form and the object, whereas their performance on the mispronunciation (MP) trials (relative to CP trials) revealed the precision of their phonological encoding: Decrements in performance with mispronunciations were taken to demonstrate accuracy in encoding (e.g., Swingley & Aslin, 2000). The two chief questions of interest were whether children who had been prefamiliarized with the novel word's form would perform better on the CP trials than the other children, indicating better learning of the linkage between the sound form and the concept, and whether prefamiliarization would result in greater mispronunciation effects, indicating more complete or confident encoding of phonological form.

Participants

Forty-four Dutch children (22 boys, 22 girls) of about 19 months of age formed the final sample. They ranged in age from 18 months 2 days to 20 months 16 days, with a mean of 18 months 24 days. Children were recruited by invitation letters sent to a list of addresses provided by the city of Nijmegen, in the Netherlands. Half of the children were assigned to the preexposure condition and half to the control condition, with equal numbers of boys and girls in each group. All children were full-term, well-baby births with Dutch as the native language. An additional 15 children were tested but not retained in the final sample because of fussiness (failure to attend on at least 10 test trials; $n = 13$), experimenter error or equipment failure ($n = 1$), or parent peeking at the visual stimuli ($n = 1$).

Overview of Stimuli and Procedure

Before arriving in the laboratory, parents completed a Dutch version of the MacArthur-Bates Communicative Development Inventory (Words and Sentences; Fenson et al., 1994), originally produced by Maryline Lejaegere for studying Flemish children and modified for the Nijmegen population by Swingley (2003). Parents were asked to indicate their children's receptive and productive vocabulary on this checklist.

Children entered the testing room with their parent, who chatted briefly with the experimenter before being seated with her child on her lap about 80 cm from the video screen. The screen was a 127-cm video projection screen (Sony KL-X9200M) that formed the back of a three-sided booth 2 m tall, 1.3 m wide, and 1.2 m deep. The speech stimuli were delivered from the loudspeakers of the video screen, and children were videotaped onto DV cassettes via a low-light camera placed about 15 cm below the monitor. Room lighting was dim. Parents were instructed that once the procedure began they were to close their eyes and orient their head downward. Initially noncompliant parents were reminded; thus, parents were unable to see the visual stimuli. Stimuli were presented using Macromedia Director software.

The experiment began with an animated story. The protagonist, a horse, wondered aloud where she might find a *tiebie* (or *droekel*; both are nonsense in Dutch). On her quest to find a *tiebie*, she encountered a cat, a duck, a field of flowers, and a fish and engaged in brief discussions of whether there might be a *tiebie* there. But the other characters professed ignorance, and the desired object was never found. This narrative (recapitulated in the Appendix) was intended as a naturalistic way to expose children to an unfamiliar word without providing any details about its meaning (and certainly without offering any information about the referent object's appearance). The novel word was used 14 times in the story.

The story phase was immediately followed by a teaching phase in which a novel image (a plush lizard or bug) was displayed alone on the screen, accompanied by prerecorded naming utterances like *Dit is een tiebie* ("This is a *tiebie*"). Children also saw the alternative novel image (the bug or lizard) accompanied by generic commentaries like *Kijk! Wat is dat?* ("Look! What is that?"). Thus, the teaching phase showed children two novel objects paired with roughly equivalent amounts of engaging verbal commentary, but only one of the objects was given a name. For half of the children (the preexposure group), this name was the one they had been exposed to in the story; for the other children (the control group), this name was entirely novel (because they had heard the story with the other word in it). In an effort to break up the possibly monotonous teaching phase, two filler trials were included midway through the teaching phase. They were of the same form as the filler trials in the test phase (described below). Thus, the teaching phase consisted of an object-naming event, a generic commentary event with the other novel object, two filler trials, another generic commentary event, and finally a second object-naming event. The novel object label was used eight times during the teaching phase.

When the teaching phase was complete, the testing phase began. It consisted of 25 trials (including the two filler trials run in the teaching phase). On each trial, two pictures were simultaneously displayed side by side for 2 s (the familiarization period). Then a speech stimulus was played, naming one of the objects. The objects remained on the screen about 3.5 s after the target word began. Trials were separated by about 0.5 s in which the screen was black.

Eight of the 25 trials were filler trials, in which two familiar objects (from a set of eight) were displayed and one was named. These objects had not been present in the story. Eight more trials showed the two novel objects and named one using the taught word in its correct pronunciation. Six trials showed the novel objects and named one using a mispronunciation. Of these MP trials, three used a subtle, one-feature mispronunciation, and three used a more conspicuous, multiple-feature mispronunciation involving adding or deleting a consonant (described in more detail below). An additional three trials displayed the two novel objects and asked for one of them using a word, *scova,* not otherwise used in the experiment. These unknown-word trials were helpful in evaluating whether children as a whole had a tendency to simply look at the taught novel object regardless of our verbal instructions. If this was true, target fixation on unknown-word trials would rival target fixation on CP and MP trials.

To help minimize repetition in the test trials, one of three 4–5-s audiovisual displays was shown between some test trials. These scenes included a fish swimming, a flower growing and being approached by a butterfly, and a kangaroo bounding across the screen.

A number of counterbalancing constraints were built in to the trial orders. Target side was equally often the left and right sides on CP trials, MP trials, and fillers. Target side was never repeated more than once on consecutive trials, and the novel word or its variants were never the target on more than two consecutive trials. Eight experimental orders were created, fully crossing the factors story (*tiebie* or *droekel*), taught word (*tiebie* or *droekel*), and referent (*lizard* or *bee*). Thus, for example, of the children in the control condition, half had heard *tiebie* in the story and were taught *droekel*, and the other half experienced the reverse.

The story animation lasted about 90 s, and the teaching and test phases together lasted about 4 min 15 s. When the procedure was finished, the parents were alerted and offered the opportunity to watch the displays.

Visual stimuli. In the story phase of the experiment, the visual display included a background image of a grassy field and cloudy sky created using a digital painting program. Other visual elements demanded by the story (a fish pond, flowers, clouds) were also hand drawn. Story characters were digital photographs of animals. During teaching and testing, the visual stimuli were digitized photographs of objects on a gray background. The novel objects used as referents for the words children were taught included a lizard (an image of a plush, bright green lizard toy with purple stripes) and a bug (an image of a plush, plump, lavender insect toy). On filler trials, children also saw images of familiar objects such as cars, cows, and ducks. Each picture measured about 28 cm wide on-screen during the teaching phase and about 21 cm wide during the testing phase (picture size was reduced slightly in testing to maintain adequate separation between pictures for eye movement coding).

Auditory stimuli. All speech stimuli were digitally recorded by a female native speaker of Dutch in a soundproofed room, sampling at 48 kHz. Her speaking rate was slow, and she spoke in a moderately infant-directed register. The sentences used for teaching the novel word *tiebie* and for commenting on the distracter picture were as follows. For the first half of the teaching phase: Target: *Dit is een tiebie. Tiebie. Een tiebie. Tiebie. Zie je de tiebie?* ("This is a tiebie. Tiebie. A tiebie. Tiebie. Do you see the tiebie?"). Distracter: *Kijk! Wat is dat? Vind je 'm mooi?* ("Look! What is that? Isn't it pretty?"). For the second half of the teaching phase: Distracter: *Kijk eens. Wat is dat? Leuk, hé?* ("Have a look. What is that? Nice, eh?"). Target: *Dit is een tiebie. Hoe noem je dat? Een tiebie! Dat is een tiebie.* ("This is a tiebie. What do you call that? A tiebie. That is a tiebie."). The same carrier phrases were used to teach the other word, *droekel*. These words were pronounced as follows: *tiebie*, [tibi]; *droekel*, [dʀukəl].

Some studies have shown that the learning of novel words that sound very similar to already familiar words (i.e., minimal pairs, or phonological neighbors) differs from the learning of more distinct words (e.g., Swingley & Aslin, in press). In the present experiments we examined the learning of words that lack phonological neighbors. No words on our Dutch MacArthur-Bates Communicative Development Inventory could be converted into *tiebie* or *droekel* by substituting, adding, or deleting a single phoneme (i.e., there were no neighbors).

Trials in the test phase included two sentences, although the analysis considered only responses to the first (as in previous studies, the second sentence was included only to help maintain children's interest). On test trials, the carrier phrase was *Waar is de* [target]? ("Where is the [target]?"). Second sentences on CP trials were *Kan je de* [target] *vinden*? ("Can you find the [target]?"), whereas second sentences on MP trials were *Kan je hem vinden?* ("Can you find it?"). For the CP targets *tiebie* ([tibi]) and *droekel* ([dʀukəl]), the relatively close, subtle mispronunciations were [kibi] and [tʀukəl]. The more distant, conspicuous mispronunciations were [kʀibi] and [tukəl]. Two levels of mispronunciation were included because children sensitive to a distant mispronunciation might not also be sensitive to a close, single-feature mispronunciation. Durations of the target words ranged from 619 to 640 ms.

On filler trials, the carrier phrase of the first sentence was *Zie je de*. . . ("Do you see the. . ."), and the targets were *auto, baby, bal, beer, eend, fiets, koe,* and *poes* ("car," "baby," "ball," "bear," "duck," "bike," "cow," and "cat"). Second sentences on these trials were either *Leuk, hé?* ("Nice, eh?") or *Vind je 'm mooi?* ("Isn't it pretty?"). Targets on filler trials averaged 515 ms in duration (SD = 86).

Coding and analysis. Videotapes of the children's faces as they looked at the screen were stamped with a timecode labeling each video frame at 40-ms intervals. These films were digitized to MPEG format and coded using custom software. Highly trained coders stepped through the MPEGs frame by frame, noting for each frame whether the child was looking at the left or right picture. This response-timing information was synchronized with stimulus-timing information generated by machine detection of tone pulses that had been aligned with speech stimulus events spliced into a stereo channel unheard by the participants. This enabled accurate estimation of the temporal relationship between children's hearing the target words and their fixation responses. Reliability was assessed by a second coder who recoded a set of eight consecutive trials from 8 randomly selected infants. Mean percent agreement was 93.9%, and mean Cohen's kappa was .909. As in previous studies (e.g., Swingley & Aslin, 2000; Swingley & Fernald, 2002), all results were computed over a short time window extending from just after the acoustic onset of the target word (360 ms after target onset) until 2.0 s after the onset.

Results

The first question of interest was whether children learned the words. This was evaluated by comparing children's target fixation proportions to the 50% that would be expected by chance.[3]

When children in the preexposed group heard the novel target word, they fixated the taught referent object 61.9% of the time (SD = 19.0), which was significantly above 50%, $t(21) = 2.93$, $p < .01$. Children not preexposed to the target word responded similarly, fixating the target 61.1% of the time (SD = 16.4), $t(21) = 3.18$, $p < .005$.

[3] Raw fixation proportions and proportions adjusted for familiarization-period looking yielded the same pattern of results.

Two further analyses were done to examine the possibility that children fixated the target object not because they recognized the word referring to it but because they found it more attractive than the distracter or because it had previously been labeled. First, if children found the target object more attractive, they should have preferred it even before it was named (i.e., in the familiarization period and during the "Where's the" phrase before the target word). This was not the case: Pre-target-word target fixation averaged 50.4% in the preexposure group and 50.8% in the control group. Second, if children looked at the referent object regardless of the speech stimuli, they should have fixated the target when the word *scova* was used. In fact, target fixation on *scova* trials was only slightly above 50%: in the preexposure group, 55.0% (SD = 31.5), $t(21) < 1$; in the control group, 55.1% (SD = 23.4), $t(21) = 1.01$, $p > .3$.

Children in the preexposure group did not perform better than children in the control group on the CP trials (61.9% versus 61.1% target fixation), $t(42) < 1$. Of the 22 children in each group, the number of children fixating the target above 50% was comparable (16 in the first group, 14 in the second). Indeed, performance on CP trials did not differ from performance on filler trials in either group: repeated measures analysis of variance (ANOVA), familiarization condition versus filler/CP, maximum $F(1, 42) = 1.3$, $p > .25$. Thus, performance on correct pronunciations was good, and early familiarization with the to-be-taught word did not appear to help children learn the link between the spoken form and its referent.

However, children in the preexposure group were much more sensitive to mispronunciations. An ANOVA with familiarization condition (preexposure or control) and word form (*tiebie* or *droekel*) as between-subjects factors and pronunciation (CP, MP-close, MP-distant) as a repeated measure revealed a significant interaction between pronunciation and familiarization, $F(2, 80) = 4.89$, $p = .01$, partial $\eta^2 = .11$, with preexposed children showing larger effects of the close MP than the control children. The ANOVA also revealed a main effect of word form, $F(1, 40) = 4.99$, $p < .05$, partial $\eta^2 = .11$, with an advantage to *tiebie*, and a marginal main effect of familiarization, $F(1, 40) = 3.41$, $p = .07$, with better performance in the preexposure condition.[4] In addition, a marginal main effect of pronunciation was found, $F(2, 80) = 2.91$, $p = .06$. No other main effects or interactions approached significance.

The interaction between pronunciation and familiarization is displayed in Figure 1. Figure 1 shows that children in both groups performed equally well when the taught word was pronounced correctly (leftmost pair of bars), but only children who had been preexposed to the novel word's form revealed sensitivity to the more subtle close MPs (viz., *kiebie* for *tiebie* and *troekel* for *droekel*).

Among the children preexposed to the taught words, correct pronunciations elicited significantly greater target fixation than either sort of mispronunciation: MP-close, $t(21) = 3.13$, $p < .001$ (one-tailed), Hedges's $g = 0.92$; MP-distant, $t(21) = 2.07$, $p = .025$ (one-tailed), $g = 0.42$. Collapsing over both types of mispronunciation yielded a mispronunciation effect of 13.2% in these children, $t(21) = 3.50$, $p = .001$ (one-tailed), $g = 0.797$. However, no effects of mispronunciation were significant in children not preexposed to the words: MP-close, $t(21) < 1.0$; MP-distant, $t(21) = 1.2$, $p > .10$; MP overall, 2.3% effect, $t(21) < 1.0$. The failure of children in the control group to show an effect of close MPs was unlikely to be due to insufficient statistical power; on the basis of the preexposure group's results, the likelihood of a Type II error was less than 10% ($\delta = 3.13$; estimate for $\alpha = .05$).[5] Rather, the lack of a mispronunciation effect in the control group was due to the nature of those children's knowledge of the novel word's form.

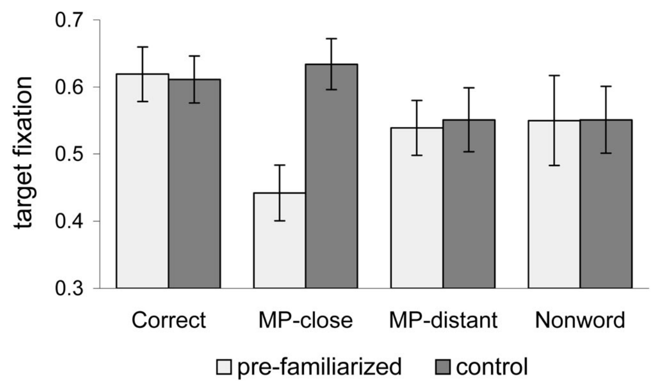

Figure 1. Children's proportion of fixation to the named picture in four conditions: Correct, when the newly taught word was correctly pronounced; MP-close, when the word was mispronounced with a single-feature change; MP-distant, when the word was mispronounced with a multiple-feature change; and Nonword, when an entirely novel nonword was produced. MP = mispronunciation; Prefamiliarized = children prefamiliarized to the taught word; Control = children in the control condition who had not received prefamiliarization. Error bars are standard errors of the mean.

A related analysis considered whether children in each group fixated the novel object either above or below chance levels when hearing its name mispronounced. Above-chance fixation to the target would be expected if the mispronounced word activated children's representation of the word enough to cause recognition; below-chance fixation would be expected if children considered the mispronunciation to be a different word with a distinct referent and inferred that the distracter picture was this referent (e.g., Halberda, 2003; Markman & Wachtel, 1988). However, children in the preexposed group responded at chance levels for both sorts of mispronunciation: MP-close, $t(21) = -1.39$, $p > .10$ (two-tailed); MP-distant, $t(21) < 1.0$; MP overall, $t(21) < 1.0$. Children not preexposed to the word fixated the target object above chance levels for close MPs, $t(21) = 3.51$, $p < .005$ (two-tailed), but were at chance for distant MPs, $t(21) = 1.08$, $p > .10$ (two-tailed). Thus, children viewing two novel objects, one with a recently taught name and one without, did not appear to consistently draw the inference that an MP (or even a wholly novel word like *scova*) must refer to the object that had never been labeled. A very similar

[4] Better learning of *tiebie* might be supposed to arise from a number of factors, including its simpler syllable structure and repeated vowel. However, in a follow-up study reported below, with the same teaching conditions, a small advantage for *droekel* was obtained. Given the inconsistency of these effects, I do not speculate further about their origins.

[5] Estimating population parameters from children the same age in Swingley and Aslin (2000) yielded a lower estimate of the likelihood of spurious null effects ($\delta = 4.80$, power $> .99$).

effect was found by Swingley and Aslin (in press). One explanation is that toddlers are most likely to use this process-of-elimination strategy (most commonly referred to as *mutual exclusivity*; Markman, 1989) when one of the objects and its name are highly familiar. Thus, children reject a spoon as a possible *dax*, favoring a novel object (e.g., Markman & Wachtel, 1988), but they do not consistently reject a "tiebie" object as a possible *scova*, perhaps because the connection between the novel word and its referent is still somewhat tenuous. This hypothesis merits further empirical attention but is not pursued here.

Although children appeared to treat the close and distant MPs differently (see Figure 1), within-subjects comparisons did not yield significant differences between the MP conditions in either group: preexposure, mean difference = 9.6% (SD = 28.9), $t(21) = 1.57$, $p > .10$; control, mean difference = 8.3% (SD = 24.3), $t(21) = 1.60$, $p > .10$. Examination of fixations during the silent familiarization period of each trial revealed a slight bias in favor of the target on distant MPs (preexposure group, 54.9%; control group, 53.4%, both *ns*), suggesting that performance on these trials might be overestimated. This is because children sometimes tend to maintain their fixation location across the trial regardless of the speech stimulus, so if children have an accidental head start before the target word, this can boost their recognition scores. Subtracting mean familiarization-period target fixation from each condition's test-period fixation proportions is a way to nullify any such biases. Analyses based on these bias-corrected data yielded the following mean scores for each group, with the preexposed group's score given first: correct pronunciation, 10.5%, 10.4%; MP-close, −3.6%, 11.1%; MP-distant, −1.0%, 1.8%; nonword, −1.0%, −2.4%. The pattern of significance in all paired comparisons using these data was identical to that in the raw percent-to-target data.

To summarize, then, children who had been familiarized to the target word before learning its meaning showed significant decrements in responding to both kinds of mispronunciation. Children without this additional exposure to the words did not show significant mispronunciation effects but hinted at sensitivity to the more severe distortions in that their target fixation fell to chance levels on the MP-distant trials.

A final series of analyses considered individual differences in performance and their potential relationship with spoken and receptive vocabulary size. This was motivated by suggestions that phonological representation develops in step with vocabulary size (e.g., Metsala & Walley, 1998). Reported comprehension vocabularies ranged from 51 to 512 words (Mdn = 190.5; M = 211, SD = 106.1), and spoken vocabularies ranged from 2 to 350 words (Mdn = 47.5; M = 69, SD = 67.1). Neither vocabulary measure was correlated with the size of children's mispronunciation effects, whether the two subject groups were considered separately or together. Although the fairly small sample sizes here preclude strong conclusions on the basis of these null results, the present lack of relationship between mispronunciation sensitivity and vocabulary size is consistent with previous studies using this method (Bailey & Plunkett, 2002; Swingley, 2003; Swingley & Aslin, 2000, 2002).

Discussion

The 22 children in the preexposure condition heard the form of a to-be-taught word 14 times in a natural story context that did not reveal the meaning of the word. They were then taught explicitly that the word named a particular novel object. Following this teaching, children's recognition of this word was tested using correct pronunciations that matched the story and teaching tokens of the word or mispronunciations that deviated from the training. These children's recognition performance was significantly worse for mispronunciations than for correct pronunciations. A separate group of 22 children in the control condition also heard a story before being taught a word, but for them the story repeated a word different from the to-be-taught word. Children in this control group were equally capable of recognizing the novel word when it was correctly pronounced but showed no sensitivity to slight mispronunciations (MP-close trials) and mixed performance on multiple-feature mispronunciations (MP-distant trials).

These results suggest three conclusions. First, although children this age can reliably differentiate phonologically similar novel words in the audiovisual habituation or switch task (Werker et al., 2002) and are acutely sensitive to mispronunciations of familiar words (Swingley & Aslin, 2000), phonological interpretation is not error free, even when the teaching utterances are specifically designed to reduce perceptual difficulties. Children in the control group heard the novel word eight times, always ending a sentence, in clearly articulated ostensive utterances and in the presence of a single novel object. Even so, at test they were apparently indifferent to whether an initial [t] was produced as [k], or [d] as [t].

Second, additional exposure appears to result in better defined phonological representations. This in itself is not a surprise, given that learning and memory would be expected to improve with repetition of the materials to be learned (e.g., Schwartz & Terrell, 1983). However, it is useful to observe that the number of repetitions that make the difference between success and failure is fairly small, at least with success defined in terms of mispronunciation sensitivity. The several dozen exposures to novel words received by children in Werker and colleagues' studies (Stager & Werker, 1997; Werker et al., 2002) are not necessary for adequate encoding (though they may be required methodologically to drive a dishabituation response). Numerical observations of this sort are important in guiding one's intuitions about children's vocabulary knowledge.

Third, children can make progress in word learning before they know what a word refers to. This point underscores the relevance of infant word-form learning to later language development. As the performance of the control group here suggests, at 1.5 years of age children are not particularly efficient in their encoding of phonological structure while learning words. This weakness is mitigated by the fact that children have a head start in building their vocabulary: Starting from at least as young as 8 months, infants extract words, or possible words, from the speech they hear (e.g., Jusczyk & Hohne, 1997), and current evidence suggests that many such words are phonologically well specified (Jusczyk & Aslin, 1995; Swingley, 2005a; Vihman et al., 2004). Thus, although word learning is generally conceptualized as the formation of connections between phonological and conceptual representations, for many words the process probably begins earlier in development with the independent acquisition of the phonological and conceptual components.

One unexpected result of this experiment was that in spite of failing to show effects of mispronunciation, children in the control group were nevertheless as good as children in the preexposure group at recognizing the novel words given correct pronunciations. One interpretation is that the control children had vague representations of words, not incorrect representations. On a binary phonological feature account, control children could be said to have had no specification of, for example, the place feature of the initial [dʀ] of *droekel*, and as a result both [dʀ] and [tʀ] were equally good matches to the stored representation. If they had had incorrect representations, overall performance in responding to correct pronunciations should have been worse in control children, but this was not the case.

This interpretation requires a distinction between the consequences of *matching* features (as when [d] in the speech signal *dog* matches a /d/ in the representation of "dog"), *mismatching* features (as when a [t] in the mispronounced *tog* conflicts with /d/ in the representation of "dog"), and *nonmismatching* features (as when a [t] or [k] in the signal neither confirms nor denies a vaguely represented consonantal onset of "tiebie").[6] Of course, matching features should be ideal cues to the word, and mismatching features should hinder recognition. Intuition suggests that the nonmismatch situation should lead to intermediate performance. Two components of the testing conditions may have prevented this result. First, the children were repeatedly asked about the same object several times, which may have led them to adopt a liberal recognition criterion when uncertain (though not when given a clear mismatch, as shown by the preexposed group). Second, the taught words had no phonological neighbors, so even if control children had scant knowledge of some features, few competing words would interfere with recognition. Note that these factors apply equally to other assumptions about phonological representation, including a more probabilistic featural model or an exemplar model. It remains an open possibility that under more difficult listening conditions (e.g., with words in sentence-medial position; Fernald, McRoberts, & Swingley, 2001) or with less predictable or acoustically less clear speech (Zangl, Klarman, Thal, Fernald, & Bates, 2005), children tested on words for which they had vague lexical representations would perform at a disadvantage.

Because this aspect of the results was unexpected, however, we undertook a second experiment with an additional group of 40 toddlers (mean age = 19 months 12 days; range = 17 months 22 days to 21 months 14 days). In this replication experiment, only CP, nonword, and filler trials were included. The MP trials were replaced with filler trials to test the possibility that hearing alternative forms of the taught words could have had an influence on their representation. For example, suppose that children in the control group started the test phase thinking that *droekel* began with [du. . .], perhaps biased by the much greater frequency of [d] than [dʀ] as a word onset in Dutch. Over the course of the test phase, hearing onsets [dʀ], [tʀ], and [t] might have led children to shift from their initial incorrect [du] to the observed vague specification, simply because the input in the test phase was consistent with multiple consonantal onsets. A more pure test of CP-trial performance would not include the MP trials. Thus, the replication experiment compared preexposed and control children on correct pronunciations of newly taught words. In all other respects, the procedure of the first experiment was maintained.

Once again, children in both groups tended to look at the target upon hearing it named: preexposed (M = 62.6%, SD = 16.2), $t(19) = 3.46$, $p < .002$; control (M = 61.3%, SD = 14.8), $t(19) = 3.43$, $p < .002$ (one-tailed). Performance in the two groups did not differ, $t(38) < 0.3$. Of the 20 children in the preexposed group, 17 fixated the target above 50% of the time; in the control group, 16 did. In addition, children did not tend to look at the "tiebie" or "droekel" picture upon hearing the nonword *scova:* preexposed (M = 52.5%, SD = 23.2); control (M = 48.4%, SD = 24.2); both ts$(19) < 1$. Thus, in both groups, children learned the connection between the spoken word and the pictured referent object, and performance did not differ by group.[7] The replication provided further evidence that under the present conditions additional exposures did not lead to improvement in recognizing the correct form, despite the fact that the correct form was likely less well specified among children not given these exposures.

How should less well-specified lexical knowledge be characterized? One possibility suggested above is a phonological account in which features are left unrepresented pending further learning. On this account, the additional familiarization provided in the preexposed group led these children to acquire the correct feature for the onsets of the words. Alternatively, if features are stored with probabilities rather than binary values, control children might have stored equal probabilities for, for example, the [t] and [k] of *tiebie*, or at least probabilities sufficiently similar for differences to remain undetected in these experiments. Here, the difference between control and preexposed children concerns the latter group's greater confidence in the correct encoding, leading to mispronunciation effects: If [k] has a much lower probability than [t], [k]-initial realizations may be considered inappropriate (or at least dubious) realizations of *tiebie*. Finally, on a lexical exemplar theory, children having heard few exposures might know that the word *tiebie* sounds like "tiebie" but not have accumulated enough tokens to have any confidence that a token with a "k" onset is substantially outside the range of variation for *tiebie*.

One would expect the failure of children to accurately or completely encode phonological details in a large number of words to make speech comprehension more difficult by increasing ambiguity. A child who knows the words *duck, cup,* and *bug* but lacks specification of the voicing or place-of-articulation features in the consonants would have no phonological basis for identifying a new word like *cut* as novel. Furthermore, even when words are not wholly ambiguous, similarity between words results in inferior recognition performance due to competitive interference (e.g., Garlock, Walley, & Metsala, 2001; Luce & Large, 2001; Luce & Pisoni, 1998; Magnuson, Tanenhaus, Aslin, & Dahan, 2003). Vague encoding renders nominally distinct words more similar to each other and should therefore inflate phonological neighborhood density and impair recognition. Thus, as their vocabularies grow, children stand to benefit from accurate phonological encoding both for speaking and for understanding language, even if vague rep-

[6] This tripartite distinction is central to the featurally underspecified lexicon (FUL) model of lexical representation and recognition (Lahiri & Reetz, 2002). However, in the FUL model, the nonmismatch situation is argued to arise from phonological underspecification of certain features even in the well-learned words of the adult lexicon and not as a consequence of incomplete learning.

[7] Analyses examining performance in relation to reported vocabulary size found that CP-trial performance was weakly correlated with vocabulary (comprehension, $r = .332$, $p < .05$; production, $r = .273$, $p = .088$). Filler trial performance was correlated with vocabulary size (comprehension vocabulary, $r = .506$, $p < .001$; production, $r = .444$, $p < .005$).

resentation of a novel word (particularly one with no neighbors) does not have a marked effect on recognition of that word.

Frequency of Lexical Exposure in Infant Development

Children in the present experiments learned a word and were tested on it within the space of several minutes. The more common situation in the child's experience involves long-term exposure to a word, perhaps with incremental accrual of semantic information over time. As a result, quantitative features of the present results (essentially, 22 repetitions being sufficient, and 8 not, for inducing single-feature mispronunciation effects) are not strictly generalizable to children's vocabularies as a whole. Children's memory for word forms may suffer degradation due to forgetting over delays; on the other hand, temporally distributed learning experience is superior to massed experience (e.g., Childers & Tomasello, 2002; Dempster, 1996; Schwartz & Terrell, 1983). It may nevertheless be useful to consider just how many words children hear on the order of 20 times to help calibrate our intuitions about the number of lexical entries children might be expected to encode in detail (or with confidence sufficient for mispronunciation sensitivity).

One way to do this is to estimate the number of words children have heard *n* times over a span of *x* days. This estimation can be done using a corpus of transcribed infant-directed speech. Of course, in considering such an analysis it is important to be mindful of the fact that transcribers bring a great deal of linguistic knowledge to bear in interpreting a recording and may "fill in" or clarify the signal in ways that are beyond infants' abilities. Nevertheless, examination of corpora is uniquely informative about word frequency.

One of the largest collections of infant-directed speech corpora, that of Brent and Siskind (2001), was used for this purpose (the available Dutch corpora are much smaller). The Brent and Siskind corpus collection contains transcriptions of ten to fifteen 75-min recordings of the speech environment of each of 14 children ranging in age from 9 to 15 months.[8] In aggregate, the 14 corpora contain about 488,000 words (tokens) of child-directed speech and represent about 207 hr of recorded time. Of this time, about one fourth contains actual child-directed speech (parents do not talk to their children nonstop). Under the assumption that children are spoken to at rates comparable to those found in the corpus for 10 hr a day, the aggregate corpus may be taken to represent (207/10) or 20.7 days' worth of speech.

Given these data describing about 3 weeks of language exposure, how many different words can one suppose children hear over 1 month or half a year? Of these words, how many occur 20 times or more? To answer these questions, the number of words (types) occurring 1, 5, 20, and 50 times in cumulative increments of the corpus were tabulated (see Figure 2, left panel). Naturally, larger and larger chunks of the corpus yielded increasing numbers of words of all frequencies, with the largest increases in words occurring only a few times. Over the approximately 3 weeks covered by the data (i.e., at 488,000 word tokens), the number of different words occurring 20 times or more was 1,450; the number occurring 50 times or more was almost 875.

Infants start learning at least some word forms before 6 months of age (Bortfeld et al., 2005), but word learning per se is usually assumed to begin just before the first birthday. Thus, the 18- and 19-month-olds in the present studies had been learning at least some word forms for several months. The right panel of Figure 2 shows a polynomial extrapolation of the original data to 7.2 million tokens, the average estimated exposure for a 10-month period. This extrapolation suggests that from 10 to 19 months of age, children may hear on the order of 5,000 words at least 20 times each and over 3,000 words at least 50 times each. Although we acknowledge once more that the present experimental data cannot fix with certainty a particular *n* instances at which children should have accurately encoded the words to which they have been exposed, it is at least plausible that children's lexicons may contain a few thousand word forms at 18 months, even if many or most of those words are not associated with any semantic content.

Note that the estimate of 3 weeks for hearing almost half a million words is an average over the 14 children. In fact, parents varied a great deal in how much they spoke (e.g., Hart & Risley, 1995). The left panel of Figure 2 relates time in weeks to the cumulative number of word tokens by showing the mean number of tokens children were estimated to experience in 1 week and also the variability in this mean. In this sample, 2 children heard less than half of the number of word tokens as the average, and a few heard half again as many as the average. This variability needs to be borne in mind when estimating relations between word-type counts and developmental time.

Conclusions

Although the literature on word learning in young children is dominated by studies of fast mapping (Diesendruck & Markson, 2001; Heibeck & Markman, 1987; Tomasello & Barton, 1994), a complete picture of the earliest stages of language acquisition requires understanding how each word in the child's vocabulary is subject to incremental changes at multiple levels of description; each word has a history with its own developmental course (Carey & Bartlett, 1978). For example, children tested on words they know may in fact have only partial semantic knowledge that will be filled in with additional exposure (e.g., Landauer & Dumais, 1997; McGregor, Friedman, Reilly, & Newman, 2002; Storkel, 2001). Pronunciation of individual words changes with experience (e.g., Macken, 1980; Schwartz, 1995; Vihman & Velleman, 2000), and the recognition of words in speech becomes faster and more reliable over time (Fernald, Perfors, & Marchman, 2005; Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998). Although children's inferences about word meaning in their initial encounters with words are often impressive (e.g., Bloom, 2000; Goodman, McDonough, & Brown, 1998; Waxman, 1999), fast mapping is not the end of the learning process, and sometimes it is not even the beginning.

What do children retain when they hear a word? Any number of circumstantial details might in principle be associated with words when they occur in conversation. By 18 months of age, children apparently consider many of these details relatively unimportant to the word's meaning, given their assumption that words refer to categories and not to whole perceptual experiences (e.g., Gelman, 2003; Markman, 1989; though see Snyder, Bates, & Bretherton, 1981, on younger children). But the fidelity with which even the linguistically relevant aspects of the conversational situation are

[8] Two additional corpora in the collection are incomplete and were excluded.

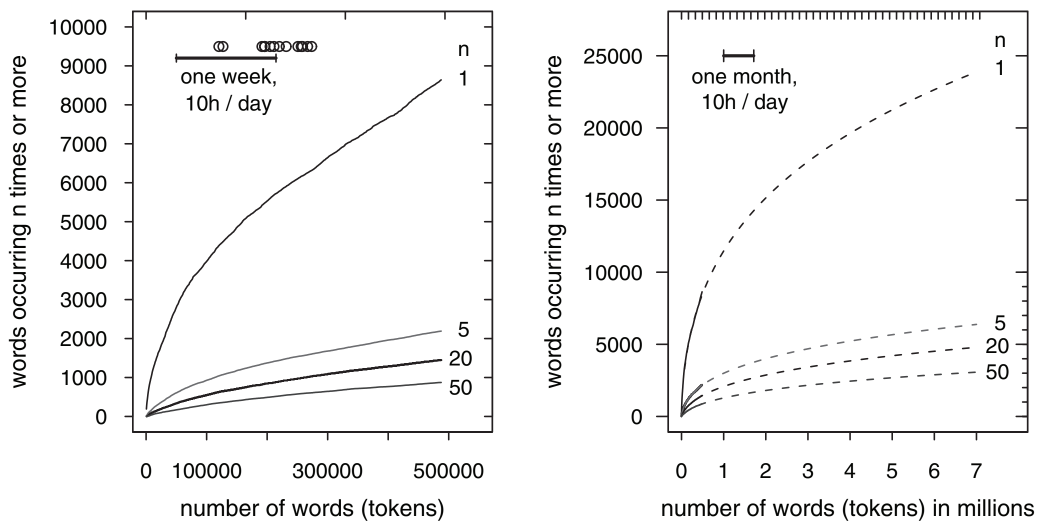

Figure 2. Word frequencies in infant-directed speech. Analyses of the Brent and Siskind (2001) speech corpus showing the total number of different words in the corpus (types) as a function of corpus size (tokens), and equivalent type counts for words that occur at frequencies of at least 5, 20, and 50. The left panel shows data from the full corpus. The line segment at the upper left gives an estimate of the mean number of word tokens heard in 1 week (165,000) on the same scale as the *x*-axis of the plot, assuming 10 hr of conversational interaction per day. Variability in the individual children contributing to this mean is shown by the circles (one circle per child). The top line of the plot's frame is demarcated in ticks corresponding to 1 average week. The right panel shows the same data (solid lines) and extrapolations thereof (dashed lines) up to about 7 million tokens, approximately 10 months' worth of exposure. The line segment at the upper left shows an estimate of the number of tokens heard in 1 month (715,000) on the scale of the *x*-axis. Ticks on the top line demarcate estimated weeks of exposure.

encoded in memory is unclear. For example, in the present studies children in the preexposure conditions heard a novel word repeatedly used as a common noun naming an object; the word usually followed an article, the protagonist spoke of searching for it, and so forth. It seems likely that this kind of information, like phonological details, could be stored in an early lexical representation that is otherwise rather lean. Experimental studies show that 1-year-olds can form categories based on the phonological contexts of novel words (Gerken, Wilson, & Lewis, 2005; Höhle, Weissenborn, Kiefer, Schulz, & Schmitz, 2004). Analyses of transcribed infant-directed speech corpora (Mintz, Newport, & Bever, 2002; Redington, Chater, & Finch, 1998; Swingley, 2005b) suggest that these computational abilities could help inform infants about word properties like form class. Thus, there is reason to believe that young children's vocabulary knowledge is substantially underestimated by measures that count a word only when children seem to know what it means. For many of children's early words, adultlike interpretation of a word's meaning may be the last step of the learning process, preceded by incremental acquisition of a well-specified phonological representation and perhaps some sense of form class and typical phrasal position.

References

Aslin, R. N., Jusczyk, P. W., & Pisoni, D. B. (1998). Speech and auditory processing during infancy: Constraints on and precursors to language. In W. Damon (Series Ed.), D. Kuhn, & R. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (5th ed., pp. 147–198). New York: Wiley.

Bailey, T. M., & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development, 17,* 1265–1282.

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

Bloom, P. (2000). *How children learn the meanings of words.* Cambridge, MA: MIT Press.

Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and me: Familiar names help launch babies into speech-stream segmentation. *Psychological Science, 16,* 298–304.

Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition, 81,* B33–B44.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Proceedings of the Stanford Child Language Conference, 15,* 17–29.

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

Dempster, F. N. (1996). Distributing and managing the conditions of encoding and practice. In E. L. Bjork & R. A. Bjork (Eds.), *Handbook of cognition and perception: Vol. 10. Memory* (pp. 317–344). New York: Academic Press.

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

Fennell, C. T. (2006). Infants of 14 months use phonetic detail in novel words embedded in naming phrases. *Proceedings of the 30th annual Boston University Conference on Language Development* (pp. 178–189). Somerville, MA: Cascadilla Press.

Fennell, C. T., & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech, 46,* 245–264.

Fennell, C. T., & Werker, J. F. (2004). Infant attention to phonetic detail:

Knowledge and familiarity effects. *Proceedings of the 28th Annual Boston University Conference on Language Development* (pp. 165–176). Somerville, MA: Cascadilla Press.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5, Serial No. 242).

Ferguson, C. A. (1964). Baby talk in six languages. *American Anthropologist, 66,* 103–114.

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development, 60,* 1497–1510.

Fernald, A., McRoberts, G. W., & Swingley, D. (2001). Infants' developing competence in recognizing and understanding words in fluent speech. In J. Weissenborn & B. Höhle (Eds.), *Approaches to bootstrapping: Phonological, lexical, syntactic, and neurophysiological aspects of early language acquisition* (Vol. I, pp. 97–123). Amsterdam: Benjamins.

Fernald, A., Perfors, A., & Marchman, V. A. (2005). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the second year. *Developmental Psychology, 42,* 98–116.

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

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20,* 104–113.

Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age of acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and Language, 44,* 468–492.

Garnica, O. K. (1973). The development of phonemic speech perception. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 215–222). New York: Academic Press.

Gelman, S. A. (2003). *The essential child.* Oxford, England: Oxford University Press.

Gerken, L. A., Wilson, R., & Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language, 32,* 249–268.

Goodman, J. C., McDonough, L., & Brown, N. B. (1998). The role of semantic context and memory in the acquisition of novel nouns. *Child Development, 69,* 1330–1344.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition, 87,* B23–B34.

Hallé, P. A., & de Boysson-Bardies, B. (1996). The format of representation of recognized words in the infants' early receptive lexicon. *Infant Behavior & Development, 19,* 463–481.

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

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

Höhle, B., Weissenborn, J., Kiefer, D., Schulz, A., & Schmitz, M. (2004). Functional elements in infants' speech processing: The role of determiners in the syntactic categorization of lexical elements. *Infancy, 5,* 341–353.

Johnson, K. (1997). The auditory/perceptual basis for speech segmentation. *OSU Working Papers in Linguistics, 50,* 101–113.

Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology, 29,* 1–23.

Jusczyk, P. W., & Hohne, E. A. (1997, September 26). Infants' memory for spoken words. *Science, 277,* 1984–1986.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5,* 831–843.

Lahiri, A., & Reetz, H. (2002). Underspecified recognition. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology 7* (pp. 637–676). Berlin, Germany: Mouton.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104,* 211–240.

Luce, P. A., & Large, N. R. (2001). Phonotactics, density, and entropy in spoken word recognition. *Language and Cognitive Processes, 16,* 565–581.

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

Macken, M. A. (1980). The child's lexical representation: The "puzzle–puddle–pickle" evidence. *Journal of Linguistics, 16,* 1–17.

Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N., & Dahan, D. (2003). The time course of spoken word learning and recognition: Studies with artificial lexicons. *Journal of Experimental Psychology: General, 132,* 202–227.

Markman, E. M. (1989). *Categorization and naming in children.* Cambridge, MA: MIT Press.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology, 20,* 121–157.

McGregor, K. K., Friedman, R. M., Reilly, R. M., & Newman, R. M. (2002). Semantic representation and naming in young children. *Journal of Speech, Language, and Hearing Research, 45,* 332–346.

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Mahwah, NJ: Erlbaum.

Mintz, T. H., Newport, E. L., & Bever, T. G. (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive Science, 26,* 393–424.

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition, 98,* 13–30.

Pater, J., Stager, C., & Werker, J. (2004). The perceptual acquisition of phonological contrasts. *Language, 80,* 384–402.

Redington, M., Chater, N., & Finch, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science, 22,* 425–469.

Schwartz, R. G. (1995). Effect of familiarity on word duration in children's speech: A preliminary investigation. *Journal of Speech and Hearing Research, 38,* 76–84.

Schwartz, R. G., & Terrell, B. Y. (1983). The role of frequency in lexical acquisition. *Journal of Child Language, 10,* 57–64.

Snyder, L. S., Bates, E., & Bretherton, I. (1981). Content and context in early lexical development. *Journal of Child Language, 8,* 565–582.

Stager, C. L., & Werker, J. F. (1997, July 24). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388,* 381–382.

Storkel, H. L. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research, 44,* 1321–1337.

Storkel, H. L. (2002). Restructuring of similarity neighbourhoods in the developing mental lexicon. *Journal of Child Language, 29,* 251–274.

Swingley, D. (2003). Phonetic detail in the developing lexicon. *Language and Speech, 46,* 265–294.

Swingley, D. (2005a). 11-month-olds' knowledge of how familiar words sound. *Developmental Science, 8,* 432–443.

Swingley, D. (2005b). Statistical clustering and the contents of the infant vocabulary. *Cognitive Psychology, 50,* 86–132.

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition, 76,* 147–166.

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

Swingley, D., & Aslin, R. N. (in press). Lexical competition in young children's word learning. *Cognitive Psychology.*

Swingley, D., & Fernald, A. (2002). Recognition of words referring to present and absent objects by 24-month-olds. *Journal of Memory and Language, 46,* 39–56.

Tomasello, M., & Barton, M. (1994). Learning words in nonostensive context. *Developmental Psychology, 30,* 639–650.

Vihman, M. M., Nakai, S., DePaolis, R. A., & Hallé, P. (2004). The role of accentual pattern in early lexical representation. *Journal of Memory and Language, 50,* 336–353.

Vihman, M. M., & Velleman, S. L. (2000). The construction of a first phonology. *Phonetica, 57,* 255–266.

Waxman, S. R. (1999). Specifying the scope of 13-month-olds' expectations for novel words. *Cognition, 70,* B35–B50.

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,* 1–30.

Werker, J. F., & Yeung, H. H. (2005). Infant speech perception bootstraps word learning. *Trends in Cognitive Sciences, 9,* 519–527.

White, K., Morgan, J. L., & Wier, L. (2004, November). *Is a dar a car? Effects of mispronunciation and context on sound-meaning mappings.* Poster session presented at the 28th annual Boston University Conference on Language Development, Boston, MA.

Woodward, A. L., & Markman, E. M. (1998). Early word learning. In W. Damon (Ed.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (pp. 371–420). New York: Wiley.

Zangl, R., Klarman, L., Thal, D., Fernald, A., & Bates, E. (2005). Dynamics of word comprehension in infancy: Developments in timing, accuracy, and resistance to acoustic degradation. *Journal of Cognition and Development, 6,* 179–208.

Appendix

Familiarization Narrative

The Dutch text of the narrated animation in which the novel word *tiebie* was presented is given, followed by a sentence-by-sentence gloss. The story for the other word, *droekel*, was identical apart from the change from *tiebie* to *droekel*. Scene changes are marked with a virgule (/).

> Hallo! Ik ben een paard. Vind je paarden leuk? Ik hoop het maar! Ik ben vandaag op zoek naar een *tiebie*. Een paar dagen geleden zag ik een *tiebie*. Hij was heel erg mooi. Ik dacht: "Ik wou dat ik een *tiebie* had." En daarom ging ik op zoek naar een *tiebie*. / Hier liep ik, op zoek naar de *tiebie*. / Ik kwam Mevrouw Kat tegen. "Mevrouw Kat, heeft u mijn *tiebie* gezien?" Ze zei: "Nee hoor, hier is geen *tiebie*." / Onderweg zag ik heel veel bloemen. Ze roken heerlijk. Maar geen *tiebie*. / Ik kwam ook Meneer Eend tegen. "Meneer Eend, heeft u een *tiebie* gezien?" Hij zei: "Nee hoor, hier is geen *tiebie*." / Als laatste zag ik Meneer Vis. "Meneer Vis, heeft u een *tiebie* gezien?" Hij maakte een visgeluid, maar had geen *tiebie* gezien. / Als snel kwamen er wolken en werd het een beetje donker. Ik had nog steeds geen *tiebie* gevonden. Maar ik bleef lopen. Ik weet niet of ik ooit een *tiebie* zal vinden! / Nou laat ik je wat plaatjes zien. Tot ziens!
>
> Hello! I am a horse. Do you like horses? I hope so! Today I'm looking for a *tiebie*. A couple of days ago I saw a *tiebie*. It was really pretty. I thought, 'I wish I had a *tiebie*.' And so I went out looking for a *tiebie*. / Here I walked, looking for the *tiebie*. / I came across Mrs. Cat. 'Mrs. Cat, have you seen my *tiebie*?' She said, 'No way, there's no *tiebie* here.' / On the way I saw lots of flowers. They smelled great. But no *tiebie*. / I also came across Mr. Duck. 'Mr. Duck, have you seen a *tiebie*?' He said, 'No way, there's no *tiebie* here.' / At last I saw Mr. Fish. 'Mr. Fish, have you seen a *tiebie*?' He made a fish noise, but had not seen a *tiebie*. / Quickly clouds came and it got a little dark. I had still found no *tiebie*. But I kept walking. I don't know if I'll ever find a *tiebie*. / Now, I'll let you see some pictures. Bye-bye!

Received August 17, 2005
Revision received July 6, 2006
Accepted July 11, 2006 ■