CHILD DEVELOPMENT

Child Development, July/August 2010, Volume 81, Number 4, Pages 1287–1303

Dogs, Bogs, Labs, and Lads: What Phonemic Generalizations Indicate About the Nature of Children's Early Word-Form Representations

Erik D. Thiessen
Carnegie Mellon University

Meagan N. Yee
Indiana University

Whereas young children accept words that differ by only a single phoneme as equivalent labels for novel objects, older children do not (J. F. Werker, C. J. Fennell, K. M. Corcoran, & C. L. Stager, 2002). In these experiments, 106 children were exposed to a training regime that has previously been found to facilitate children's use of phonemic contrasts (E. D. Thiessen, 2007). The results indicate that the effect of this training is limited to contexts that are highly similar to children's initial experience with the phonemic contrast, suggesting that early word-form representations are not composed of entirely abstract units such as phonemes or features. Instead, these results are consistent with theories suggesting that children's early word-form representations retain contextual and perceptual features associated with children's prior experience with words.

From early in life, children are able to remember words; they recognize words after familiarization (e.g., Jusczyk & Aslin, 1995) and look at the appropriate visual referent of spoken words (e.g., Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998; Tincofff & Jusczyk, 1999). But speech presents a representational challenge because of its variability. The same word can be produced with different rates, fundamental frequencies, amplitude, and formant structure both within and across speakers. Learners must represent words in a sufficiently flexible manner that two different utterances of the same word can be mapped to the same meaning. But at the same time, representations must be conservative enough that learners can determine when they are hearing a novel word (as when a child who knows *dog* hears *bog* for the first time), as opposed to a novel production of a familiar word. This is likely to be especially challenging for infants and young children, whose knowledge of the phonological system of their native language is in a state of flux (e.g., Werker & Tees, 1984).

One approach to this representational problem suggests that speech is mapped on to abstract, invariant representations (e.g., Nearey, 1997). For example, in traditional linguistic theory (e.g., Chomsky & Halle, 1968) representations of word forms in the mental lexicon are constructed from abstract, invariant phonological components such as phonemes or features (e.g., Nearey, 2001; Prince & Smolensky, 1997). These components are identical across different surface realizations: the same features are activated by the [d] in *dog* and the [d] in *lad*. According to this approach, listeners must factor out the surface variability of speech to access the abstract units that compose the word. There have been several proposals as to how this might be accomplished, including perceptual normalization accounts (e.g., Shankweiler, Strange, & Verbrugge, 1977) and articulatory theories (e.g., Browman & Goldstein, 1986, 1992; Repp, 1982). These proposals, and others, share the assumption that surface variability, such as speaker-specific indexical properties, is noise that must be discarded before listeners can identify abstract, phonological units of representation.

An alternative approach suggests that the surface variability of speech is not discarded but instead remains part of the long-term representation of word forms from early in development (e.g., Werker & Curtin, 2005). This approach is consistent

---

This research was funded by a grant from the National Science Foundation (BCS-0642415) to the first author. We thank Ashley Episcopo, Teresa Pegors, and many other research assistants for their help in running participants. Additionally, we thank Lori Holt, Katharine Graf Estes, and Janet Werker and two anonymous referees for helpful discussion and useful comments on previous versions of this manuscript, as well as the many parents who were willing to participate in the research.

Correspondence concerning this article should be addressed to Erik D. Thiessen, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213. Electronic mail may be sent to thiessen@andrew.cmu.edu.

0009-3920/2010/8104-0018

with demonstrations that listeners recognize recurring words more easily when they are produced by the same speaker, indicating that indexical information is encoded in representations (e.g., Bradlow, Nygaard, & Pisoni, 1999; Goldinger, 1996). One way listeners might encode surface variability is to store multiple exemplars of words rather than a single abstract and canonical version (e.g., Goldinger, 1998; Jusczyk, 1993). Such information may be sufficient for word recognition alone, or it may be incorporated in parallel with more abstract representations (e.g., Hanson, 1977; Remez, Fellowes, & Rubin, 1997). Even very young learners may have access to both kinds of representations. Research on categorical perception (e.g., Eimas, Siqueland, Jusczyk, & Vigorito, 1971) and vowel change detection in the face of speaker differences (Kuhl, 1983) suggest that infants can respond to acoustic differences that reflect linguistically relevant distinctions, such as phonemic categories, in the face of surface variability. At the same time, however, infants appear to encode speaker-specific properties of their experience with words (Houston & Jusczyk, 2003), and respond to acoustic differences between exemplars of a single phonemic category (McMurray & Aslin, 2005).

In the adult state, an abstract approach and approaches that incorporate surface variability in representations make similar predictions. For example, connectionist models can learn, through exposure to variable input, to respond to the same phonemic distinctions as models relying on more abstract representations (e.g., Vallabha, McClelland, Pons, Werker, & Amano, 2007). However, these two kinds of representational approaches lead to very different predictions about the rate of learning and generalization. In a theory where generalization occurs automatically at the level of abstract phonemes, a child who responds to [d]—as distinguished from [t]—has mapped the spoken realization of those phonemes onto representations abstracted from their acoustic idiosyncracies. Because these representations are abstract, they should generalize to a variety of novel acoustic instances. A learner who responds to the contrast followed by the vowel [i] should respond to the contrast in the same way when it is followed by the vowel [a]; even though the acoustic realization of [d] is different in *da* than *di*, the underlying representation is identical. Similarly, a contrast available in syllable-initial position should be available in syllable-medial or syllable-final position, insofar as the contrast can be mapped onto identical representations.

These kinds of abstract approaches predict very rapid patterns of developmental acquisition. Shvachkin (1948/1973)—working from an acoustic framework—argued that once children master the use of a particular phoneme, it should be available in every context in which the native language employs that phoneme. More recent theories have made similar predictions. For example, Lindblom, Diehl, Park, and Salvi (2008), whose work emphasizes a more articulatory approach, predict that ''**once a target** [place feature] **has been learned in one context, it can be immediately re-used in other contexts**.'' (bolding in original). This powerful generalization is one of the strengths of an abstract representational system (e.g., Berent, Marcus, Shimron, & Gafos, 2002), as it can potentially explain the rapidity, and consistency across a variety of differing environments, with which children acquire language (for discussion, see Lindblom, Diehl, Park, & Salvi, in press). By contrast, a representational system that encodes multiple aspects of the surface variability of a contrast should lead to less inclusive generalization in early learning. Because the acoustic realization of a phonemic contrast is different across different word positions and vocalic contexts, a contrast that is productive in one setting may not be productive in all settings. Instead, generalization on this kind of account should be based, at least in part, upon the similarity of the current instance to prior experience (e.g., McClelland & Elman, 1986). Further, the frequency with which a learner experiences a contrast in a particular context should be important; if the [d]–[t] contrast is more common in word-initial position than in word-final position (e.g., Zamuner, 2009), a learner might initially respond to contrast more successfully in word-initial position (e.g., Schafer & Mareschal, 2001; Storkel, 2001).

Note, however, that theories proposing an abstract representational system need not predict invariant responses to every instance of a phoneme. For example, in word recognition, different realizations of the phoneme may be more or less effective in activating the phonemic representation, as is the case when an ambiguous articulation is heard (e.g., Norris & McQueen, 2008). Similarly, in phonotactic learning, a particular phoneme may occur in some word positions or syllabic contexts, but not in others (e.g., Hayes, 2004). More generally, it is well known that distinctions at a perceptual level need not correspond with distinctions at a phonological level (e.g., Dietrich, Swingley, & Werker, 2007; Pallier, Colome, & Sebastian-Galles, 2001; Stager & Werker, 1997). Indeed, even in early childhood, there is evidence that children's sensitivity is graded as a function of the tasks in which they are tested. Children will use a distinction in a word-recognition task—as when distinguishing between familiar words like *ball* and *doll* (e.g., Fennell & Werker, 2003; Swingley & Aslin, 2000)—that they fail to use in the context of associating a novel label with a novel object (Pater, Stager, & Werker, 2004). These complexities make it particularly important to explore generalization in the context of word–object association. Simply because infants generalize in a perceptual task (e.g., Maye, Weiss, & Aslin, 2008) does not necessarily indicate that they will do so when representing word forms. Conversely, the fact that learners generalize conservatively when learning phonotactic regularities (e.g., Prince & Tesar, 2004) does not require that they will do so when learning words.

To distinguish theories where generalization occurs at the level of abstract phonemes, or context-free features (e.g., Dinnsen, O'Connor, & Gierut, 2001; Nearey, 2001), from perceptually based accounts (e.g., Houston & Jusczyk, 2003; Ryalls & Pisoni, 1997; Singh, 2008), it is critical to assess the course of learning in a task relevant to lexical development. Learning should be rapid and highly generalizable if children's representations are abstract, slower and more conservative for representations that incorporate subphonemic and indexical information (cf. McClelland & Patterson, 2002). One important component of word learning is the ability to associate labels with objects, an ability that can be assessed experimentally (e.g., Stager & Werker, 1997). Such assessments indicate that young children have difficulty differentiating between novel labels that differ by only a single phoneme (such as *bih* and *dih*), although they can learn to do so successfully (Thiessen, 2007; Yoshida, Fennell, Swingley, & Werker, 2009). In this series of experiments, we will assess how widely children generalize their learning about phonemic differences in a word–object association task.

Using the difference between words that differ by a single phoneme (i.e., minimal pair words such as *bog* and *dog*) in a word–object association task is difficult for children in the first half of their 2nd year of life (e.g., Shvachkin, 1948/1973; Stager & Werker, 1997). Upon learning a label for a novel object (such as *bin*), children of around 14 months of age accept a wide variety of minimal pairs of that word (e.g., *pin* and *din*) as labels for the novel object (Pater et al., 2004). This is not due to an inability to hear the phonemic distinction. When tested on the same contrasts with known words (such as *ball* and *doll*), children are responsive to the identical contrast (e.g., Fennell & Werker, 2003; Swingley & Aslin, 2002). Instead, children appear to have difficulty using phonemic contrasts in the context of associating novel labels with novel objects. The ability to use phonemic contrasts to differentiate novel labels improves with development. By 17 months, children are beginning to behave more like adults: After they have learned a label for a novel object, they reject alternative labels that differ on one or more phonemes. At 14 and 17 months, the ability to do so is related to vocabulary size; by 20 months, this is no longer the case, perhaps because by this age virtually all children have exceeded the threshold of vocabulary necessary to succeed in the task (Werker et al., 2002).

The relation between vocabulary and performance on the switch task posited by Werker et al. (2002) suggests that linguistic experience influences children's use of phonemic contrasts in when learning words. This would explain why children with larger vocabularies were more successful in the task. Consistent with this hypothesis, computational modeling indicates that experience with words may affect the way that children respond to phonemic contrasts (Schafer & Mareschal, 2001), and both 17- and 20-month-olds know far more words than 14-month-olds (Fenson et al., 2002). In particular, children in this age range know relatively few minimal pair words (e.g., Caselli et al., 1995; Charles-Luce & Luce, 1995). For example, Swingley and Aslin (2007) found that, in a sample of 18-month-old Dutch-learning children, over two thirds of the words in children's vocabularies had no minimal pair neighbors. The majority of the words with which children become familiar between 14 and 20 months provide examples of phonemes distributed in different lexical contexts: [d] in *doggy* and *daddy*, [t] in *teddy* and *tiger*.

Thiessen (2007) demonstrated that exposure to phonemes in different lexical contexts ([d] in *dawbow* and [t] in *tawgoo*) facilitates children's use of the phonemic contrast in a word–object association task. This may be related to the phenomenon of acquired distinctiveness: When two similar stimuli, A and B, are paired with two highly discriminable associates X and Y (forming compound stimuli AX and BY), the distinction between A and B is heightened (e.g., Hall, 1991). Children's familiarity with an increasing number of words in the 2nd year of life, very few of which are minimal pairs, may provide evidence that different phonemes are paired with different lexical forms in the language and facilitate use of the distinction between them

(Thiessen, 2007). This would explain the observed correlation between vocabulary size and use of phonemes in word–object association tasks (Werker et al., 2002); children with larger vocabularies have more evidence about the distribution of phonemes in their language. From the perspective of exemplar-based theories, these results can be explained in terms of variability among exemplars facilitating generalization (e.g., Hintzman, 1986; Rost & MacMurray, 2009). From the perspective of theories emphasizing more abstract representational systems, or parallel encoding of acoustic and abstract information (e.g., Remez et al., 1997), success on the task might indicate a shift to more abstract encoding. For example, Werker and Curtin (2005) suggest that while children's representations contain rich acoustic detail, children must weight the phonemic aspects of these representations more heavily to succeed in word-learning tasks relying on a single phonemic distinction.

As such, while the results of Thiessen (2007) indicated that exposure to phonemes in different lexical contexts improves children's abilities to use the contrast between phonemes, those experiments were ambiguous with respect to whether children learned about abstract, readily generalized contrasts (i.e., phonemes or features). The current series of experiments systematically assessed how widely infants generalize a contrast (such as [d] vs. [t]) after they have been exposed to stimuli that should facilitate their use of the contrast in a word–object association task. If children encode speech abstractly, generalization should be rapid and pervasive: Once a child has identified the [d]–[t] contrast in one context, it should be available in many (or even all) contexts (e.g., Lindblom et al., 2008; Shvachkin, 1948/1973). By contrast, if children encode subphonemic or indexical information, generalization may be much more limited, especially early in learning (e.g., McClelland & Patterson, 2002). Children should initially be most likely to use the [d]–[t] contrast in a context they have previously experienced, and less likely to use it in contexts that differ from the original context. Conversely, training in a different context should allow children to succeed in contexts where they have previously failed, as is demonstrated in Experiment 3.

Experiment 1

Fifteen-month-olds fail to respond to differences in minimal pair labels in a word–object association task (e.g., Pater et al., 2004; Shvachkin, 1948/1973). Thiessen (2007) found that exposing children to words in which phonemes occurred in clearly distinct lexical contexts—as [d] and [t] in *dawbow* and *tawgoo*—facilitates their use of the distinction between minimal pairs. However, these results are inconclusive with respect to the degree of specificity or abstraction of children's learning. Children were tested with the syllables *daw* and *taw* after exposure to the words *dawbow* and *tawgoo*. Therefore, children could potentially have learned about a feature contrast (voicing), a phonemic contrast ([d]–[t]), or a contrast in a specific syllabic context (*daw* vs. *taw*). If children learned about voicing, this knowledge should generalize to a wide variety of contexts. By contrast, if infants learned about the specific syllables, then their generalization should be limited. Similarly, most prior experiments testing children's use of phonemic contrasts expose participants to phonemes in only one particular context, making it difficult to assess how generalizable children's representation of these contrasts might be (e.g., Stager & Werker, 1997).

Our goal in this experiment was to assess how widely children generalize the contrasts they use in word–object association tasks. To do so, we exposed children to the same training regime used in Thiessen (2007): Children learned the names of three objects. The labels of two of these objects provided evidence of phonemes ([d] and [t]) occurring in distinct lexical contexts: *dawbow* and *tawgoo*. The third object was labeled *yad*. After learning the names of these three objects, infants were presented with test trials in which the third object was paired with its original label (*yad*), or a novel minimal pair label (*yat*). Experience with *dawbow* and *tawgoo* enables children to differentiate between *daw* and *taw* test trials in this procedure (Thiessen, 2007). If children have learned about a phonemic contrast, or a featural contrast (voicing), this learning should generalize to a novel word position (from [d]-[t]aw to ya[d]-[t]). By contrast, if children's learning is more specific, and tied to the particular exemplars with which children have prior experience, then exposure to *dawbow* and *tawgoo* would not be as informative with respect to *yad* as it is with respect to *daw*, and children may fail to use the contrast between [d] and [t] in word-final position.

Method

Participants

Participants were 16 toddlers (half female; 13 Caucasian, 2 Asian, and 1 African American) between the ages of 14.5 and 15.5 months

(*M* = 14.77). To obtain data from these 16 infants, it was necessary to test 24. The additional 8 participants were excluded for the following reasons: fussing or crying (7), experimental error (1). According to parental report, all children were free of ear infection at the time of testing, and reported no history of hearing problems. Children were recruited from the Pittsburgh area via mailing. The participants primarily came from middle- or upper-middle-class socioeconomic status (SES) families.

Stimuli

Participants were exposed to three novel objects, each paired with a unique label. The three objects were identical to those used by Stager and Werker (1997) and Thiessen (2007). Each object was animated and moved against a black background. Each of the three labels was produced by the SoftVoice computer synthesizer, set to a female voice, and a monotonic production with a fundamental frequency of 220 Hz.

The first object was paired with the label *yad* [jaed] or the label *yat* [jaet] (here and elsewhere, brackets indicate IPA transcription). Each infant heard only one of these labels—either *yad* or *yat*—during the habituation phase. The second object was paired with the label *dawbow* [dabo], and the third object was paired with the label *tawgoo* [tagu]. The labels *dawbow* and *tawgoo* were the identical synthesized recordings as those used in Thiessen (2007), and the labels *yad* and *yat* were synthesized to have matching pitch, amplitude, and vocal quality. Following Thiessen's original experiment, there was only a single token of each synthesized label used during the habituation and test phases. Each label was repeated for as long as the object remained on screen, with pauses of 1.4 s between repetitions.

Procedure

This experiment used a variant of the habituation procedure employed by Stager and Werker (1997) procedure, and identical to that used in Thiessen (2007). Participants, seated on a parent's lap in a sound-attenuated room, controlled the duration of the presentation of the stimuli by the length of their gaze at a central monitor. To eliminate bias, parents wore headphones, and the experimenters sat in an adjacent room, blind to the nature of the stimulus being presented. The experimenter coded the duration of the child's looking time online, using a Macintosh running the Habit program for OS X (Cohen, Atkison, & Chaput, 2004). After the child reached the habituation criterion (looking time less than 50% of the average of the first three trials), six test trials were presented.

The child controlled the duration of stimulus presentation, during both the habituation and test phases, by gazing at a 32′ video monitor 150 cm in front of their seated position. The child's attention was attracted to the monitor by a colorful video of Winnie the Pooh coupled with a recorded verbal encouragement. Once the experimenter determined the child's attention was fixated on the monitor, stimulus presentation was initiated. An object then appeared on the screen and the speakers adjacent to the monitor began to repeat the label associated with that object. The stimulus presentation continued until the child looked away for more than 1.5 s, or until the child had gazed at the monitor for 20 s (the maximum time allowed per trial). The video of Winnie the Pooh appeared at the end of each trial to recapture the child's attention.

The experiment began with two familiarization trials, in which participants saw a nonsense object paired with the word *neem* (a recording from a trained female speaker). The familiarization trials were used to help participants become accustomed to the pairing of audio and visual stimuli. Once these trials were finished, the habituation phase began.

During the habituation trials, one of the three novel objects appeared on screen, and was paired with a unique label (Object 1 was always paired with *yad* or *yat*; Object 2 was always paired with *dawbow*, and Object 3 was always paired with *tawgoo*). The objects were presented in blocks of three trials, and the order of presentation was random within those blocks. As in Thiessen (2007), each participant was required to complete at least two habituation blocks, or their data were excluded (all participants reached this threshold). Each object moved onscreen while the label was repeated with pauses of 1.4 s between each repetition. The object and the label were presented until the child looked away from the monitor for 1.5 s, at which point the attention-getting stimulus reappeared. Looking times to each trial were calculated in real time, and the habituation trials continued in random order until the child met the habituation criterion: Average looking time for three consecutive trials that fell below 50% of their looking time to the first three habituation trials.

Once the child met the habituation criterion, the six test trials began. In all test trials, the child saw Object 1 on the monitor. There were two kinds of test trials: same and switch trials. In the same trials, the children heard the syllable that they had previously heard paired with Object 1. For half of the infants, this was *yad*; for the other half of the infants, this was *yat*. In the switch trials, Object 1 was paired with a minimal pair of the label infants had heard in the habituation phase. For infants exposed to the object paired with *yad*, the switch trial paired the object with the novel label *yat*. The opposite was true for infants exposed to *yat* during the habituation phase.

Same and switch trials alternated, and the nature of the initial test trial was counterbalanced across participants. As in the habituation trials, the object stayed on the screen, and the label continued to repeat, for as long as the participant continued to look at the monitor.

Results and Discussion

On average, children habituated in 8.9 trials. Although children habituated in fewer trials to *yad* ($M = 7.5$) than to *yat* ($M = 10.2$) during the habituation phase, the difference was only marginally significant, $t(14) = 1.9$, $p = .08$, Cohen's $d = 0.55$. More importantly, children's preference for same versus switch test trials did not differ as a function of which item they heard during the habituation phase, $t(14) = 1.5$, $p = .15$, Cohen's $d = 0.41$. Therefore, for all subsequent analyses, participants were grouped together, regardless of which item they heard during the habituation phase. During the habituation phase, infants spent an average of 96.5 s ($SD = 28.6$) looking to the objects during the habituation phase. Thus, on average infants accumulated approximately 30 s of looking to the object labeled *yad/yat*.

Our primary question was whether infants in the test phase responded differentially to trials in which the novel object was paired with its original label (same trials), versus trials in which the object was paired with a minimal pair of that label (switch trials). As illustrated in Figure 1, children did not respond differentially to same and switch trials. A two-tailed t test (all t tests reported are two-tailed) indicated that the difference in looking times between same and switch trials was not significant, $t(15) < 1$, *ns*, Cohen's $d = 0.04$.

These results contrast sharply with those of Thiessen (2007). In that experiment, exposure to *dawbow* and *tawgoo* lead children to respond differentially to objects paired with *daw* versus *taw*. Here, after exposure to *dawbow* and *tawgoo*, children responded equivalently *yad* and *yat*. One possibility is that the sample size in this experiment was too small to yield a significant result; the sample size in Thiessen was 24 infants, compared to only 16 in this experiment. A power analysis indicated that for an effect the size in the Thiessen paper (approximated by a Cohen's d of .75, given the reported difference of 1.9 s between groups and a standard deviation of 2.4 s), the probability of identifying a significant result with the current sample size of 16 is only 54%. A sample size of 30 would yield a greater than 80% probability of finding an effect. Thus, in Experiment 2 we utilized a larger sample size. Another possibility is that children are unable to hear the difference between [d] and [t] in word-final position. However, the results of Experiment 3 (below) indicate that children can respond to this contrast in word-final position.

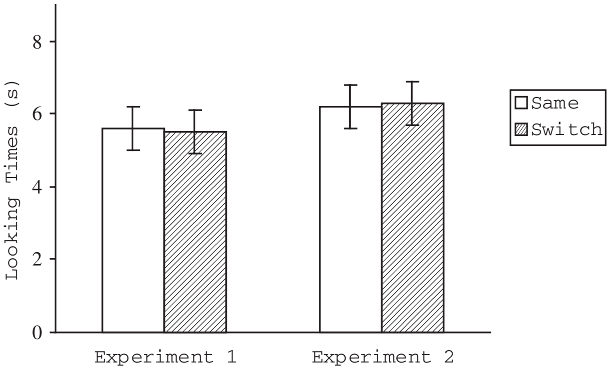

Figure 1. Children's looking time to same and switch trials after exposure to *Dawbow* and *Tawgoo*.
Note. In Experiment 1, same and switch trials are *yad* versus *yat*; in Experiment 2, they are *dee* versus *tee*. Error bars indicate ± standard error.

There is, though, a theoretical possibility consistent with the lack of a significant difference found in this experiment. What children learn from experiencing speech sounds in distinct lexical contexts (as with the [d] and [t] sounds in *dawbow* and *tawgoo*) does not lead, at least initially, to generalizations about abstract units such as features or phonemes. Instead, children may be learning about more specific, contextually bound units. That is, children's learning may be limited to situations that share a high degree of similarity to their initial experience. For example, exposure to *dawbow* and *tawgoo* may only facilitate use of [d] and [t] in word-initial position, or even be limited to a particular syllabic context. Whereas children fail to use the [d]–[t] contrast in word-final position, they may succeed if tested with contexts in which the [d]–[t] contrast is realized in a way that is more similar to its occurrence in *dawbow* and *tawgoo*.

Experiment 2

Writing and phonemic systems use the same symbols to indicate a phoneme in different word positions. However, the acoustic realization of phonemes differs in as a function of context. For example, the phoneme [t] is aspirated in *tear*, while it is unaspirated in *steer*. The [t] in *steer* sounds, when the preceding fricative is removed, much like a [d] (e.g., Maye, Werker, & Gerken, 2002). Similarly, phonemes are articulated quite differently in word-initial position than in word-final position, and these articulatory differences lead to differences in the acoustic realization of the phoneme (Stevens, 1998). If children are representing speech sounds in a nonabstract manner, the acoustic differences between word-initial and word-final position may make it difficult for children to generalize from information about a phonemic contrast across word positions; experience with *dawbow* and *tawgoo* may not be readily applicable to *yad* and *yat*.

However, testing generalization by presenting phonemes in word-final position may be especially challenging for young children. The ends of words appear to be less salient than word-onset positions for children (e.g., Clark, 1991), and word-initial information may be especially salient given its role in word recognition (e.g., Marslen-Wilson, 1987; Rodd, 2004). Given children's potential difficulty with word-final position, asking children to generalize from word-initial to word-final position may be an overly conservative test of generalization. In Experiment 2, we tested generalization from one word-initial setting to a second word-initial setting, in a different syllabic context. As in Experiment 1, children were exposed to three novel object–label pairings: objects labeled *dawbow, tawgoo*, and *dee*. According to Thiessen (2007), experience with *dawbow* and *tawgoo* facilitates children's use of the contrast between *daw* and *taw*. If children are learning about a relatively abstract contrast, they should also be facilitated in their use of the contrast between *dee* and *tee*.

Method

Participants

Participants were 30 toddlers (half female; 26 Caucasian and 4 Asian) between the ages of 14.5 and 15.5 months (M = 15.22). To obtain data from these 30 children, it was necessary to test 37. The additional 7 participants were excluded for the following reasons: fussing or crying (5), experimental error (2). According to parental report, all children were free of ear infection at the time of testing, and reported no history of hearing problems. Children were recruited from the Pittsburgh area via mailing. The participants primarily came from middle- or upper-middle-class SES families.

Stimuli

The computer-animated objects were identical to those used in Experiment 1, as were the words *dawbow* and *tawgoo*. The only difference between the stimuli used in Experiment 1, and the stimuli used in this experiment, was that Object 1 was paired with the labels *dee* [di] or *tee* [ti] (each infant heard only one during the habituation phase). These labels replaced *yad* and *yat* from Experiment 1, and were synthesized using the same acoustic parameters as the stimuli in that experiment. As in Experiment 1, there was only a single repeated token of each label.

Procedure

The procedure was identical to that of Experiment 1. After the habituation phase, infants were presented with same and switch trials. On same trials, Object 1 was paired with the same label as it was during the habituation phase. On switch trials, it was paired with a minimal pair of that label. Thus, for infants who heard *dee* during the habituation phase, the switch trial presented the word *tee*.

Results and Discussion

On average, children exposed to *dee* (in addition to *dawbow* and *tawgoo*) habituated in 13.4 trials, while children exposed to *tee* (in addition to *dawbow* and *tawgoo*) habituated in 6.9 trials. This difference was significant, $t(28) = 2.7$, $p < .05$, Cohen's $d = 1.02$. Infants exposed to *dee* spent an average of 128.1 s (SD = 75.4) looking to the objects during the habituation phase. Infants exposed to *tee* spent an average of 79.4 s (SD = 35.4) looking to the objects during the habituation phase. This difference was also significant, $t(28) = 2.3$, $p < .05$, Cohen's $d = 0.83$. It is unclear why children found one set of habituation trials more interesting than the other. It may be indicative of a preference for the speech synthesizer's production of voiced items over voiceless items, or a difference in children's familiarity with these particular syllables in their natural environment. However, any possible item preference did not influence children's performance on the test items: There was no significant difference in children's preference for same versus switch trials as a function of their exposure to *dee* or *tee*, $t(28) < 1$, *ns*, Cohen's $d = 0.06$.

Given that the differences in rate of habituation suggested that children may have performed differently as a function of exposure, we analyzed the group data separately, using a 2 (habituation: *dee* or *tee*) × 2 (test trial: same vs. switch) analysis of variance (ANOVA). There was no main effect of habituation, $F(1, 28) < 1$, *n.s.*, $\eta_p^2 = .01$. Nor was there any interaction between habituation and test item, $F(1, 28) < 1$, *ns*, $\eta_p^2 = .01$.

Most importantly, there was no difference in children's looking time to same or switch trials $F(1, 28) < 1$, *n.s.*, $\eta_p^2 = .01$. Figure 1 illustrates the average looking time to same and switch trials, averaged across both groups of infants (habituated to *dee*, or habituated to *tee*).

These results suggest that exposure to *dawbow* and *tawgoo* had little, if any, effect on children's use of the [d] and [t] contrast in a novel syllabic context. Whereas *dawbow* and *tawgoo* facilitated children's use of the contrast in the same (*daw–taw*) syllabic context (Thiessen, 2007), children continued to treat minimal pair words differentiated by the contrast interchangeably, when the contrast was embedded in a novel syllable. This indicates that what children may not readily generalize at the level of abstract phonemes, or context-free features. Instead, children's learning appears to be strongly tied to specific, previously experienced, contexts.

Experiment 3

The results of Experiment 1 and 2, taken together, suggest that children are not generalizing at the level of abstract units such as phonemes or features. Rather, their use of speech sounds to indicate contrasts in meaning appears to be contextually bound. If so, then it should be easiest to facilitate children's use of a phonemic contrast in a particular context by providing them distributional evidence about that phoneme in an identical context. To test this hypothesis, children in Experiment 3 were tested using the *yad–yat* pairing from Experiment 1. However, children received distributional information about the [d]–[t] contrast in word-final position (the words *boeyad* vs. *gooyat*). This experience with the contrast in word-final position should facilitate its use in *yad–yat*, unlike the experience with the phonemes in word-initial position in Experiment 1.

To increase the ecological validity of the comparison between word-initial and word-final position, Experiment 3 used natural speech, as opposed to synthesized speech; this necessitated a replication of the data from Experiment 1 with the new, natural stimuli. As such, there were three conditions in this experiment.

Single Object

In this condition, children were habituated to a single object, paired with the label *yad*. Following the results of Stager and Werker (1997), children should fail to respond differentially to same (*yad*) and switch (*yat*) trials.

Word Initial

This condition replicates Experiment 1. In addition to *yad*, children were habituated to two objects labeled *dawbow* and *tawgoo*, providing information about the distribution of [d] and [t] in word-initial condition. If the results of natural speech parallel the results with synthesized speech, children should also fail to respond differentially to same and switch trials.

Word Final

In this condition, children were habituated to objects labeled *yad*, *boeyad* [bojaed], and *gooyat* [gujaet]. This condition provides evidence that the [d]–[t] contrast occurs in different distributions in word-final position, and thus should be the condition in which children are most successful at responding differentially to the contrast in word-final position.

Method

Participants

Participants were 60 toddlers (half female; 51 Caucasian, 4 Asian, 4 African American, and 1 Hispanic) between the ages of 15.0 and 16.5 months ($M = 15.4$). Twenty were assigned to the single-object condition, 20 were assigned to the word-initial condition, and 20 were assigned to the word-final condition. To obtain data from these 60 participants, it was necessary to test 73. The additional 13 infants were excluded for the following reasons: fussing or crying (8), experimental error (2), parental interference (2), and looking times averaging < 2 s during test trials (1). According to parental report, all children were free of ear infection at the time of testing, and reported no history of hearing problems. Children were recruited from the Pittsburgh area via mailing. The participants primarily came from middle- or upper-middle-class SES families.

Stimuli

The computer-animated objects were identical to those used in Experiment 1. However, the audio stimuli were recorded from utterances produced by a female native speaker of English, and digitally edited to match the amplitude of the stimuli used in Experiments 1 and 2. The test items (*yad* and *yat*) were produced by the same speaker. As in prior experiments, only a single token of each label was used during the habituation and test phases.

Procedure

All participants saw Object 1 paired with the nonsense word *yad*. In the single-object condition, this was the only word–object pairing to which children were habituated. In the word-initial condition, children saw two additional objects, labeled *dawbow* and *tawgoo*. This replicates Experiment 1, providing distributional information about the [d]–[t] contrast in word-initial position.

The third condition was the word-final condition, in which children were provided with distributional information about the [d]–[t] contrast in word-final position. In this condition, children were habituated to Object 1 paired with *yad*, and two additional objects paired with *boeyad* and *gooyat*.

The habituation procedure used was identical to that of Experiment 1 and 2. After habituation, all infants were tested on same and switch trials, as in Experiment 1. On same trials, Object 1 was paired with the familiar label *yad*; on switch trials, it was paired with the novel label *yat*.

Results and Discussion

Participants in the single-word condition habituated in fewer trials ($M = 6.8$) than participants in either the word-initial ($M = 10.2$) or word-final condition ($M = 9.8$). Similarly, participants in the single-word condition spent less overall time looking during the habituation phase ($M = 58.3$ s, $SD = 28.5$) than infants in either the word-initial ($M = 112.1$ s, $SD = 47.6$) or word-final condition (104.6, $SD = 39.1$). This is unsurprising, as participants in the word-initial and word-final conditions were exposed to three word-object pairings during the habituation phase, whereas participants in the single-object condition were only exposed to one. These results mimic those of Thiessen (2007), where children habituated more quickly to a single object than multiple objects.

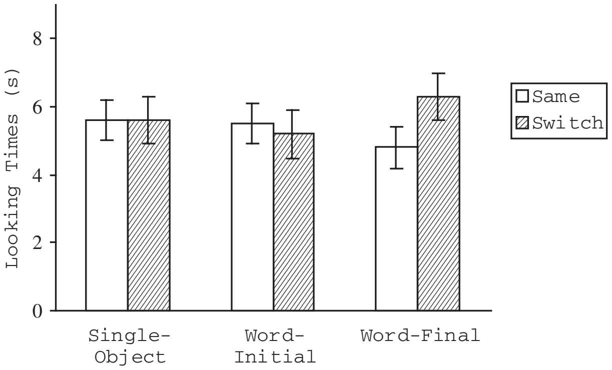

Figure 2. Children's looking time to same and switch trials in the three conditions of Experiment 3.
Note. The difference between same and switch trials was significant only in the word-final condition. Error bars indicate ± standard error.

As seen in Figure 2, participants in the single-object condition responded equivalently to same and switch trials, $t(19) < 1$, *ns*, Cohen's $d = 0.01$. Children's lack of differentiation between object labels differing by only a single phoneme replicates Stager and Werker's (1997) original experiment, and many other demonstrations that children at this age have difficulty responding to single-phoneme differences in word–object association tasks (e.g., Pater et al., 2004; Shvachkin, 1948/1973; Werker et al., 2002).

In the word-initial condition (in which infants were exposed to *dawbow* and *tawgoo*, in addition to *yad*), children also failed to respond differentially to same and switch trials, $t(19) < 1$, *ns*, Cohen's $d = 0.14$. Children's looking times to same and switch trials are illustrated in Figure 2; their failure to respond differentially to the test trials replicates Experiment 1.

Most importantly, as seen in Figure 2, children in the word-final condition responded differentially to same and switch trials. This difference was significant, $t(19) = 2.4$, $p < .05$, Cohen's $d = 0.6$. Unlike experience with *dawbow* and *tawgoo* (the word-initial condition), exposure to *boeyad* and *gooyat* facilitated children's response to the difference between *yad* and *yat*.

To more thoroughly explore the difference between the word-initial and word-final condition, we performed a 2 (condition) × 2 (test trial type) ANOVA. There was no main effect of condition, $F(1, 38) < 1$, *ns*, $\eta_p^2 = .01$. This suggests that looking times were equivalent across the two conditions. There was also no main effect of test item, $F(1, 38) = 1.2$, $p = .27$, $\eta_p^2 = .03$. This lack of effect is not surprising, as participants did not demonstrate a consistent preference for items across conditions. Most importantly, however, there was an interaction between condition and item, $F(1, 38) = 5.1$, $p < .05$, $\eta_p^2 = .12$. This interaction indicates that toddlers' preference in the word-final condition significantly differs from their lack of preference in the word-initial prediction. These results support the hypothesis that children's performance is better when they receive distributional information in the same context as the phonemes are presented in the test trials.

The difference between the word-initial and word-final conditions is informative about the nature of children's early generalization of speech sounds. In both conditions, children received distributional evidence that the phonemes [d] and [t] occur in different lexical forms; this kind of distributional evidence has previously been found to facilitate children's use of phonemic contrasts (Thiessen, 2007). However, only distributional information in the same syllable and word position facilitated children's ability to use the contrast at test. Children in the word-initial condition, who received distributional information about the phonemes in different syllables and word positions, performed like children in the single-object condition who received no distributional information at all. This suggests that children's generalization is limited, at least initially, to contexts that are highly similar to their prior experience.

One potential confound needs to be addressed: There is a possibility that children respond differently in the two conditions, not because of the context of the distributional information they received but because they receive different amounts of exposure to the stimuli. Morgan, Duran, and Layton (2005) have suggested that the ability to dishabituate in experiments of this type is mediated by amount of exposure; when children are overexposed to a particular sound, they become less sensitive to it. Alternatively, children may have received less exposure to the *yad*-object in the word-initial condition and thus not learned the pairings between sound and object. To assess these possibilities, we performed a two-tailed t test comparing length of exposure during the habituation phase for children in the word-initial condition ($M = 112.1$ s, $SD = 47.6$) and children in the word-final condition (104.6, $SD = 39.1$). There was no significant difference, $t(38) < 1$, *ns*, Cohen's $d = 0.17$. This suggests that length of exposure does not explain why children differentiate between same and switch trials in the word-final condition, and fail to do so in the word-initial condition. Instead, this analysis reinforces the hypothesis that the critical difference in these conditions relates to the similarity between the distributional information children receive during testing and the context in which phonemic distinctions occur during test trials.

As such, these results are inconsistent with accounts in which children's early representations of speech are organized in such a way as to allow automatic generalization at the level of abstract phonemes. If phonemic representations were entirely abstract (e.g., bundles of features), children's use of those phonemes should generalize widely (e.g., Berent et al., 2002; Lindblom et al., 2008; Shvachkin, 1948/1973). Instead, the current results suggest that children are storing some amount of contextual (i.e., not abstracted) detail about their prior experience with phonemes. This detail may be related to word position, syllabic context, or some other acoustic aspect of the specific realization of the previously experienced phoneme. These details persist and lead to performance that is better when the context matches training than when it does not. While the precise nature of children's representation is still ambiguous (see the General Discussion), the fact that children perform better in the word-final condition than in the word-initial condition indicates that children's representations of early word forms are rich with contextual detail.

General Discussion

To master their native language, children must learn which sound contrasts indicate a difference in meaning and which do not. While children make remarkable progress in that regard during their 1st year of life (e.g., Werker & Tees, 1984), they still have some difficulty using these distinctions in a word–object association task into the 2nd year (e.g., Stager & Werker, 1997). By 17 months, children are more successful at making use of at least some of these single phoneme differences (Werker et al., 2002), although they may not be proficient with every contrast present in their native language until later (Shvachkin, 1948/1973). While a number of explanations have been proposed for this developmental difference—including increases in capacity (Stager & Werker, 1997) and understanding of referential intent (e.g., Fennell, 2008)—the current results provide support for accounts suggesting that older children are successful because they have more information about the distribution of phonemes in their native language (e.g., Thiessen, 2007).

Children's vocabularies grow dramatically during the 2nd year of life (e.g., Fenson et al., 2002). During this time, most of the words children learn provide evidence of phonemes occurring in different lexical contexts (like /b/ and [d] in *baby* and *diaper*). Young children know comparatively few words in which two different phonemes occur in identical contexts (minimal pair words like *ball* and *doll*), far fewer, both absolutely and as a relative proportion of their vocabulary, than adults (e.g., Coady & Aslin, 2003). As demonstrated by Experiment 3, and also by Thiessen (2007), experiencing phonemes in different lexical contexts helps children to use the phonemic distinction in word–object association tasks. Further, children with larger vocabularies, and thus more evidence that phonemes are distributed in different words, are more successful using phonemic distinctions in word–object association tasks (Werker et al., 2002). These results converge to suggest that vocabulary growth plays an important role in children's use of phonemic contrasts by providing children with distributional information.

The current results are novel in that they indicate that children's generalization of the distributional information they receive is somewhat limited. Children are more likely to take advantage of distributional information about a phonemic contrast when that contrast is tested in the same context as the original distributional information. While these results do not rule out the possibility that some aspects of children's representation of word forms are abstract (e.g., featural), they indicate that children must be storing some contextual detail about their prior experience with phonemes. This is inconsistent with theories that argue that representation of speech is primarily or completely abstract. For example, most featural theories of representation argue that phonemes can be reduced to bundles of abstract features that are invariant across word position or syllabic context (e.g., Chomsky & Halle, 1968; Cole & Scott, 1974; Eulitz & Lahiri, 2004). It is a common assumption that children's representations of speech can be characterized by the same kinds of abstract representations (e.g., Dinnsen et al., 2001). The context specificity of the current results is inconsistent with abstract representations like features or phonemes, although it may be the case that abstract representations and contextual acoustic detail are processed separately and in parallel (e.g., Hanson, 1977; Remez et al., 1997). One possibility is that abstract representations, such as phonemes, can incorporate contextual detail such as the syllabic context in which the phoneme occurred in ways that limit the generalizations learners make about those phonemes in other contexts (e.g., Prince & Tesar, 2004).

Indeed, these results, and others like them (e.g., Houston & Jusczyk, 2003) are consistent with accounts positing that representations of speech are not composed of abstract, phonologically pure components (e.g., Goldinger, 1996). While phonemes and features are useful descriptively, the evidence for the psychological reality of phonemes has been criticized on a number of grounds (e.g., Lotto & Holt, 2003). Many phenomena once thought to be decisive evidence in favor of abstract representations, such as context effects, have now been demonstrated to occur in the absence of phonemic knowledge (e.g., Lotto, Kluender, & Holt, 1997; Mann, 1986). Similarly, intraphonemic variation, which should be discarded on abstract representational accounts, clearly plays a role in adult speech perception (e.g., Volaitis & Miller, 1992). Although both adults and infants appear to respond to phonological categories in ways that are consistent with abstract representations (e.g., Eimas et al., 1971), it may be the case that nonabstract systems could accomplish the same feats (Bybee & McClelland, 2005). For example, when infants generalize on the basis of a feature such as voicing (e.g., Maye et al., 2008), this generalization may be due to acoustic similarities among voiced consonants, rather than a commonality in the abstract featural representation + voicing. However, caution is due in interpreting the current results; evidence that children's representations are not completely abstract is not logically sufficient to rule out some degree of abstraction.

Some authors have suggested that the degree of abstraction in children's representation varies by the task in which they are engaged, or their developmental state (e.g., Shvachkin, 1948/1973). Child phonologists have proposed a discontinuity between the types of representations used in perceptual tasks (such as speech discrimination) and the representations used for word learning and language more broadly (e.g., Barton, 1980; Brown & Matthews, 1997; Ferguson & Farwell, 1979). One possibility that this raises is that children's representations in the current tasks looked relatively concrete because they were not treating the task as truly linguistic, but rather as a perceptual association task. On this kind of account, had the training been embedded in a more naturalistic setting, children may have engaged a different set of representations and generalized their phonemic usage more widely. By contrast, Werker et al. (2002) have proposed that the underlying representations used by infants in both perceptual and linguistic tasks are identical. Due to the fact that word-learning and perceptual tasks typically differ in their difficulty (word-learning tasks are more difficult because they require forming an association between words and objects), children show differing degrees of sensitivity to phonemic variation in the two kinds of tasks. While the current data do not differentiate between these two accounts, other work indicates that lessening the difficulty of word-learning tasks—whether by using familiar objects (e.g., Fennell, 2006), or prefamiliarized labels (e.g., Swingley, 2007)—leads infants to perform more successfully in word–object association tasks. This pattern of results is more consistent with theories positing a continuous representational system than theories positing discontinuous representational systems.

This should not be taken as evidence that there are no developmental differences in children's representations over time. A number of theorists have argued that more abstract representations emerge over time (e.g., Best, Tyler, Gooding, Orlando, & Quann, 2009; Dietrich et al., 2007). Werker and Curtin (2005), for example, have proposed that children's representational system encodes subphonemic and indexical information throughout the course of development. As children gain more experience with their native language, they come to attend more to phonemic aspects of these representations than to more idiosyncratic factors. This may explain why older children are more successful in using phonemic detail in word–object association tasks (e.g., Werker et al., 2002). The emergence of less context-specific phonemic distinctions may occur as a function of children's experience with these contrasts in more varied contexts (e.g., Rost & MacMurray, 2009). In these experiments, children failed to generalize a phonemic contrast after experiencing it in a single vocalic or syllabic context. However, generalization may be possible after experiencing the contrast in multiple contexts, as suggested by models of generalization in memory (e.g., Hintzman, 1986). As a contrast is experienced in multiple different contexts, the information that is consistent—the contrast itself—is emphasized, while the variable contextual information is less supported in memory. The emergence of behavior governed by (seemingly) abstract units from contextually detailed representations is consistent with a variety of formal models.

However, before the predictions of such models can be fully evaluated, it is critical to understand the nature of the contextual detail encoded in children's representations. While the current results demonstrate that children's representations of speech incorporate some kind of contextual detail about their prior experience with phonemes, the nature of that contextual information is still uncertain. Two (potentially complementary) possibilities have been suggested in prior theories of linguistic development: context at a grain size above the phoneme, such as syllables, and more granular acoustic context, such as coarticulatory information. The syllable is a universal feature of human linguistic systems; additionally, syllables are both perceptually salient and relatively limited in number. Infants appear to be sensitive to the presence and identity of syllables in speech (e.g., Bertoncini, Bijeljic-Babic, Jusczyk, Kennedy, & Mehler, 1988; Jusczyk & Derrah, 1987). Because of these facts, a number of theorists have argued that syllables may play an important role in children's early representations of speech (e.g., Mehler, Dupoux, & Segui, 1990). On these kinds of accounts, the syllable is an undifferentiated unit, not composed of combinatorially distinct subcomponents. This lack of subcomponents (phonemes or features) renders a syllabic account consistent with the current results. When children are exposed to *dawbow* and *tawgoo*, they are learning about the syllables *daw* and *taw*, not the phonemes [d] and [t] or the features [−voice] and [+voice]. Therefore, there is little reason to expect generalization to syllable contrasts like *yad/yat* or *dee/tee*. This syllable-level account is consistent with theories of phonotactic learning in which generalization is conditioned upon the syllabic context in which a particular phoneme has been previously encountered (e.g., Prince & Tesar, 2004).

An alternative, although not mutually exclusive, possibility is that different instances of particular phonemes might be differentiated by fine-grained acoustic differences in the realization of the phoneme itself. Both infants and adults are sensitive to intraphonemic differences: differences in the articulation of two exemplars of the same phonemic category (e.g., McMurray & Aslin, 2005; Volaitis &

Miller, 1992). Representation of these intraphonemic differences is often assumed in exemplar memory models of speech representation (e.g., Goldinger, 1998). Similarly, the WRAPSA model of speech perception (Jusczyk, 1993, 1997) proposes that the developmentally earliest form of representation is a spectral and temporal code derived immediately from the acoustic signal itself. The fact that infants recognize words more easily after a delay when they are spoken by the same speaker than when they are produced by a novel speaker (Houston & Jusczyk, 2003; Singh, 2008) is consistent with approaches emphasizing a fine-grained acoustic basis of representation, because it indicates that infants' representation of the word forms preserves idiosyncratic acoustic indexical information such as rate of speech. If young children are representing intraphonemic acoustic details, generalization should also be more limited than is the case with an abstract phonological system. This is because different speech sounds are realized quite differently in different contexts, due to coarticulatory processes. The realization of [d] in *yad* is quite different than the realization in *dawbow*. By contrast, the [d] in *boeyad* is nearly identical to the [d] in *yad*, making distributional information from *boeyad* much easier to apply to *yad* than distributional information from *dawbow*.

Subsequent experimentation will be necessary to assess the precise nature of the contextual detail present in young children's representations of word forms. The current results demonstrate that children fail to generalize to a segment that shares a feature (like voicing) in a very different syllabic position, or followed by a very different vowel. But children might be successful if they were tested on the same feature—in the same syllabic and vocalic context—at a different place of articulation. For example, children might successfully use the distinction between *baw* and *paw* after exposure to *dawbow* and *tawgoo* (we thank an anonymous reviewer for this suggestion). Alternatively, children may generalize to a different syllable as long as the vocalic context and syllabic position are the same. That is, after exposure to *dawbow* and *tawgoo*, children may use the distinction between *dock* and *talk*. If children are representing intraphonemic acoustic variation, they should be successful in this training scenario (a variant of which we are currently running in our lab), since the phonemes [d] and [t] are articulated in the same way in *dawbow* and *dock* and *tawgoo* and *talk*, occurring in the same syllabic position and vocalic context. By contrast, if children are representing syllabically, they may fail in this training regime, since *dock* and *talk* are novel syllables.

The current results do not distinguish between a syllabic and coarticulatory account of the contextual information stored in young children's representations of early word forms (for a discussion, see Jusczyk, 1993; Coady & Aslin, 2004). But on either account, consistent with the current results, children's representations are not characterized solely by abstract phonemic or featural components. This leads to a different view of the developmental processes between the age when children fail to use phonemic differences in word-learning tasks (Stager & Werker, 1997), and the age when they succeed (Werker et al., 2002). If children are representing words using featural or phonemic representations, the change between failure and success is somewhat mysterious, as children are already using appropriate units. One way to resolve this dilemma is to propose that children's early representations of speech are holistic and underspecified (e.g., Walley, 1993). For example, a word like *dog* might, early in lexical development, be represented only by the single feature +alveolar. However, recent research on comprehension is largely inconsistent with an underspecification account (although see Kager, Van der Feest, Fikkert, Kerkhoff, and Zamuner, 2007, for production data consistent with underspecification). Young children are highly sensitive to even slight mispronunciations of known words (e.g., Swingley & Aslin, 2000, 2002, 2007), and even some novel words (Mani & Plunkett, 2008).

Rather than underspecification, these results suggest that young children's early word learning may fail to take advantage of phonemic differences due to the richness of the input and children's representations (cf. Werker & Curtin, 2005). There are a multitude of acoustic differences to which children are sensitive (e.g., Houston & Jusczyk, 2003; McMurray & Aslin, 2005). Some of these acoustic differences indicate differences in meaning, but many do not. Children may need to receive information from their linguistic environment about which of these differences are informative in word-learning tasks before they begin to use them productively. For example, hearing examples of two different phonemes distributed in different word forms helps children subsequently use that phonemic contrast when using novel words (Thiessen, 2007). The current results are consistent with accounts suggesting that children's representations are rich and multidimensional, insofar as they indicate children's representations are not entirely abstracted but instead incorporate contextual detail. To the extent that children's representations incorporate variability, children may need to receive evidence about which facets of that variability are productive in particular tasks.

This kind of contextual sensitivity, coupled with a learning account, leads to novel predictions about the developmental time course of children's use of different phonemic contrasts. Abstract representational accounts assert that the development in children's use of phonemic contrasts should be sudden (e.g., Lindblom et al., 2008; Nazzi & Bertoncini, 2003); once a child has learned to use a featural contrast (e.g., voicing) in one situation, that contrast can be widely used. By contrast, the current results indicate limited generalization and predict a more graded development in children's use of phonemic contrasts in word learning. More common contrasts should begin to be used earlier than less common contrasts. Similarly, a sound that occurs disproportionately often in word-initial position should be productive in that position earlier than it is productive in other word positions (cf. Schafer & Mareschal, 2001). While there is currently little data to assess these predictions (although see Shvachkin, 1948/1973; Thiessen & Yee, 2008), the current results provide an empirical demonstration that children's use of a phonemic contrast in one context does not immediately generalize to other contexts.

Note, however, that the present findings indicating conservative generalization do not mean that young learners are incapable of generalizing their experience with speech sounds. In particular, prior research suggests that variability in the input facilitates generalization (e.g., Gomez, 2002; Lively, Logan, & Pisoni, 1993). Maye et al. (2008) found that infants who learned to discriminate a featural contrast were able to generalize that learning to a novel contrast varying on the same feature; in that experiment, infants were exposed to multiple exemplars of the featural contrast. Participants in these experiments received only a single exemplar of each to-be-learned label. Young children may well be able to generalize in word–object association tasks if they receive training that more robustly supports generalization beyond specific contexts, such as experiencing a phonemic contrast (e.g., [d]–[t]) in multiple syllabic contexts. This would be consistent with both prior work (e.g., Maye et al., 2002; Rost & MacMurray, 2009) and the current experiments, indicating that infants are sensitive to the distribution of acoustic details in the input and that this distribution plays an important role in the development of children's linguistic abilities.

References

Barton, D. (1980). Phonemic perception in children. In G. Yeni-Komshian, J. Kavanagh, & C. Ferguson (Eds.), *Phology: Vol. 2 perception* (pp. 97–116). New York: Academic.

Berent, I., Marcus, G. F., Shimron, J., & Gafos, A. I. (2002). The scope of linguistic generalizations: Evidence from Hebrew word formation. *Cognition*, *83*, 113–139.

Bertoncini, J., Bijeljic-Babic, R., Jusczyk, P. W., Kennedy, L. J., & Mehler, J. (1988). An investigation of young infants' perceptual representation of speech sounds. *Journal of Experimental Psychology: General*, *117*, 21–33.

Best, C. T., Tyler, M. D., Gooding, T. N., Orlando, C. B., & Quann, C. A. (2009). Development of phonological constancy: Toddlers' perception of native- and Jamaican-accented words. *Psychological Science*, *20*, 539–542.

Bradlow, A. R., Nygaard, L. C., & Pisoni, D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Perception and Psychophysics*, *61*, 206–219.

Browman, C. P., & Goldstein, L. (1986). Towards an articulatory phonology. In C. Ewen & J. Anderson (Eds.), *Phonology yearbook* (pp. 219–252). Cambridge, UK: Cambridge University Press.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, *49*, 155–180.

Brown, C., & Matthews, J. (1997). The role of feature geometry in the development of phonemic contrasts. In S. Hannahs & M. Young-Scholten (Eds.), *Focus on phonological acquisition* (pp. 67–112). Amsterdam: Benjamins.

Bybee, J., & McClelland, J. L. (2005). Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *The Linguistic Review*, *22*, 381–410.

Caselli, M. C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L., et al. (1995). A cross-linguistic study of early lexical development. *Cognitive Development*, *10*, 159–199.

Charles-Luce, J., & Luce, P. A. (1995). An examination of similarity neighborhoods in young children's receptive vocabularies. *Journal of Child Language*, *22*, 727–735.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.

Clark, E. V. (1991). Acquisitional principles in lexical development. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 31–71). Cambridge, UK: Cambridge University Press.

Coady, J. A., & Aslin, R. N. (2003). Phonological neighbourhoods in the developing lexicon. *Journal of Child Language*, *30*, 441–469.

Coady, J. A., & Aslin, R. N. (2004). Young children's sensitivity to probabilistic phonotactics in the developing lexicon. *Journal of Experimental Child Psychology*, *89*, 183–213.

Cohen, L. B., Atkison, D. J., & Chaput, H. H. (2004). *Habit X: A new program for obtaining and organizing data in* *infant perception and cognition studies (Version 1.0)*. Austin: University of Texas.

Cole, R. A., & Scott, B. (1974). Toward a theory of speech perception. *Psychological Review*, *81*, 348–374.

Dietrich, C., Swingley, D., & Werker, J. F. (2007). Native language governs interpretation of salient speech sound differences at 18 months. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 454–464.

Dinnsen, D. A., O'Connor, K. M., & Gierut, J. A. (2001). The puzzle-puddle-pickle problem and the Duke-of-York gambit in acquisition. *Journal of Linguistics*, *37*, 503–525.

Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*, 303–306.

Eulitz, C., & Lahiri, A. (2004). Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition. *Journal of Cognitive Neuroscience*, *16*, 577–583.

Fennell, C. T. (2006, June). *Object familiarity affects 14-month-old infants' use of phonetic detail in novel words*. Paper presented at the biennial International Conference on Infant Studies, Kyoto, Japan.

Fennell, C. T. (2008, March). *Does coarticulatory information aid use of phonetic detail at 14 months?* Paper presented at the International Conference on Infant Studies.

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.

Fenson, F., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., et al. (2002). *MacArthur Communicative Development Inventories: User's guide and technical manual*. Baltimore: Paul H. Brookes.

Ferguson, C. A., & Farwell, C. B. (1979). Words and sounds in early language acquisition. *Language*, *51*, 419–439.

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

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1166–1183.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*, 251–279.

Gomez, R. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*, 431–436.

Hall, G. (1991). *Perceptual and associative learning*. Oxford, UK: Clarendon.

Hanson, V. L. (1977). Within-category discriminations in speech perception. *Perception and Psychophysics*, *21*, 423–430.

Hayes, B. (2004). Phonological acquisition in Optimality Theory: The early stages. In R. Kager, J. Pater, & W. Zonneveld (Eds.), *Fixing priorities: Constraints in phonological acquisition* (pp. 158–203). Cambridge, UK: Cambridge University Press.

Hintzman, D. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*, 411–428.

Houston, D. M., & Jusczyk, P. W. (2003). Infants' long-term memory for the sound patterns of words and voices. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 1143–1154.

Jusczyk, P. W. (1993). From general to language-specific capacities: The WRAPSA model of how speech perception develops. *Journal of Phonetics*, *21*, 3–28.

Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.

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., & Derrah, C. (1987). Representation of speech sounds by young infants. *Developmental Psychology*, *23*, 648–654.

Kager, R., Van der Feest, S. V. H., Fikkert, P., Kerkhoff, A. O., & Zamuner, T. S. (2007). Representations of [voice]: Evidence from acquisition. In J. van der Weijer & E. J. van der Torre (Eds.), *Voicing in Dutch* (pp. 41–80). Amsterdam: John Benjamins.

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

Lindblom, B., Diehl, R., Park, S., & Salvi, G. (2008, June). *(Re)use of place features in voiced stop systems: Role of phonetic constraints*. Paper presented at the annual FONETIK conference, Gothenburg, Sweden.

Lindblom, B., Diehl, R., Park, S., & Salvi, G. (in press). Sound systems are shaped by their users: The recombination of phonetic substance. In N. Clements & R. Ridouane (Eds.), *Where do features come from? The nature and sources of phonological primitives*. Elsevier: North-Holland.

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/: II. The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, *94*, 1242–1255.

Lotto, A. J., & Holt, L. L. (2003). The illusion of the phoneme. In S. J. Billings, J. P. Boyle, & A. M. Griffith (Eds.), *Chicago Linguistic Society: Vol. 35. The panels* (pp. 191–204). Chicago: Chicago Linguistic Society.

Lotto, A. J., Kluender, K. R., & Holt, L. L. (1997). Perceptual compensation for coarticulation by Japanese quail (*Coturnix coturnix japonica*). *Journal of the Acoustical Society of America*, *102*, 1134–1140.

Mani, N., & Plunkett, K. (2008). Fourteen-month-olds pay attention to vowels in novel words. *Developmental Science*, *11*, 53–59.

Mann, V. A. (1986). Distinguishing universal and language-dependent levels of speech perception: Evidence from Japanese listeners' perception of English "l" and "r." *Cognition*, *24*, 169–196.

Marslen-Wilson, W. (1987). Functional parallelism in spoken word recognition. *Cognition*, *25*, 71–102.

Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science*, *11*, 122–134.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*, B101–B111.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86.

McClelland, J. L., & Patterson, K. (2002). Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Science*, *6*, 465–472.

McMurray, B., & Aslin, R. N. (2005). Infants are sensitive to within-category variation in speech perception. *Cognition*, *95*, B15–B26.

Mehler, J., Dupoux, E., & Segui, J. (1990). Constraining models of lexical access: The onset of word recognition. In G. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 236–262). Cambridge, MA: MIT Press.

Morgan, J., Duran, P., & Layton, D. (2005, November). *A U-shaped pattern in infants' lexical representations?* Paper presented at the annual Boston University Conference on Language Development.

Nazzi, T., & Bertoncini, J. (2003). Before and after the vocabulary spurt: Two modes of word acquisition? *Developmental Science*, *6*, 136–142.

Nearey, T. M. (1997). Speech perception as pattern recognition. *Journal of the Acoustical Society of America*, *101*, 381–394.

Nearey, T. M. (2001). Phoneme-like units in speech perception. *Language and Cognitive Processes*, *16*, 673–681.

Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, *115*, 357–395.

Pallier, C., Colome, A., & Sebastian-Galles, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science*, *12*, 445–449.

Pater, J., Stager, C. L., & Werker, J. F. (2004). The lexical acquisition of phonological contrasts. *Language*, *80*, 361–379.

Prince, A., & Smolensky, P. (1997). Optimality: From neural networks to Universal Grammar. *Science*, *275*, 1604–1610.

Prince, A., & Tesar, B. (2004). Learning phonotactic distributions. In R. Kager, J. Pater, & W. Zonneveld (Eds.), *Fixing priorities: Constraints in phonological acquisition* (pp. 245–291). Cambridge, UK: Cambridge University Press.

Remez, R. E., Fellowes, J. M., & Rubin, P. E. (1997). Talker identification based on phonetic information. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 651–666.

Repp, B. (1982). Phonetic trading relations and context effects: New evidence for a phonetic mode of perception. *Psychological Bulletin*, *92*, 81–110.

Rodd, J. M. (2004). When do leotards get their spots? Semantic activation of lexical neighbors in word recognitions. *Psychonomic Bulletin & Review*, *11*, 434–439.

Rost, G., & MacMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science*, *12*, 339–349.

Ryalls, B. O., & Pisoni, D. B. (1997). The effect of talker variability on word recognition in preschool children. *Developmental Psychology*, *3*, 441–452.

Schafer, G., & Mareschal, D. (2001). Modeling infant speech sound discrimination using simple associative networks. *Infancy*, *2*, 7–28.

Shankweiler, D., Strange, W., & Verbrugge, R. (1977). Speech and the problem of perceptual constancy. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: Toward an ecological psychology* (pp. 315–345). Hillsdale, NJ: Erlbaum.

Shvachkin, N. Kh. (1973). The development of phonemic speech perception in early childhood. In C. Ferguson & D. Slobin (Eds.), *Studies of child language development* (pp. 91–127). New York: Holt, Rinehart, & Winston. (Original work published 1948)

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

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

Stevens, K. N. (1998). *Acoustic phonetics*. Cambridge, MA: MIT Press.

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

Swingley, D. (2007). Lexical exposure and word-form encoding in 1.5-year-olds. *Developmental Psychology*, *43*, 454–464.

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. (2007). Lexical competition in young children's word learning. *Cognitive Psychology*, *54*, 99–132.

Thiessen, E. D. (2007). The effect of distributional information on children's use of phonemic contrasts. *Journal of Memory and Language*, *56*, 16–34.

Thiessen, E. D., & Yee, M. N. (2008, March). *Evidence that variability in lexical frames affects children's use of phonemic contrasts*. Paper presented at the International Conference on Infant Studies.

Tincofff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, *10*, 172–175.

Vallabha, G. K., McClelland, J. L., Pons, F., Werker, J. F., & Amano, S. (2007). Unsupervised learning of vowel categories from infant-directed speech. *Proceedings of* *the National Academy of Sciences of the United States of America*, *104*, 13273–13278.

Volaitis, L. E., & Miller, J. L. (1992). Phonetic prototypes: Influence of place of articulation and speaking rate on the internal structure of voicing categories. *Journal of the Acoustical Society of America*, *92*, 723–735.

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

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

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., & Tees, R. C. (1984). Cross language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*, 49–63.

Yoshida, K., Fennell, C., Swingley, D., & Werker, J. F. (2009). 14-month-olds learn similar-sounding words. *Developmental Science*, *12*, 412–418.

Zamuner, T. S. (2009). The structure and nature of phonological neighbourhoods in children's early lexicons. *Journal of Child Language*, *36*, 3–21.