Journal of Experimental Psychology:
Human Perception and Performance
2003, Vol. 29, No. 6, 1143–1154

0096-1523/03/$12.00 DOI: 10.1037/0096-1523.29.6.1143

Infants' Long-Term Memory for the Sound Patterns of Words and Voices

Derek M. Houston
Indiana University

Peter W. Jusczyk
Johns Hopkins University

Infants' long-term memory for the phonological patterns of words versus the indexical properties of talkers' voices was examined in 3 experiments using the Headturn Preference Procedure (D. G. Kemler Nelson et al., 1995). Infants were familiarized with repetitions of 2 words and tested on the next day for their orientation times to 4 passages—2 of which included the familiarized words. At 7.5 months of age, infants oriented longer to passages containing familiarized words when these were produced by the original talker. At 7.5 and 10.5 months of age, infants did not recognize words in passages produced by a novel female talker. In contrast, 7.5-month-olds demonstrated word recognition in both talker conditions when presented with passages produced by both the original and the novel talker. The findings suggest that talker-specific information can prime infants' memory for words and facilitate word recognition across talkers.

Understanding spoken language involves segmenting fluent speech into words and accessing representations of those words from memory. Native language users can recognize familiar words effortlessly, even when those words have not been encountered in days, months, or longer. During the 2nd year of life, normally developing children hit their word-learning stride, rapidly encoding many words a day (Bloom, 1975; Clark, 1973). To learn words, language learners must link sound patterns to their appropriate meanings. Although many studies have focused on how children attach the appropriate meanings to words (e.g., see A. L. Woodward & Markman, 1997, for a review), considerably less is known about what they first encode into memory regarding the sound patterns of words and when they begin this encoding.

In a typical language environment, the vast majority of words that infants hear occur in the context of fluent speech rather than in isolation (van de Weijer, 1998; J. Z. Woodward & Aslin, 1990). Hence, before they can encode the sound patterns of words into memory, infants must first segment them from fluent speech. Pioneering work by Jusczyk and Aslin (1995) showed that by 7.5 months of age, infants display some ability to segment monosyllabic words from fluent speech. Infants who were familiarized with two words (repeated in citation form) and then tested on four passages—two of which contained the familiarized words—listened significantly longer to the passages with the familiarized words than to control passages. These findings suggest that the infants stored some information about the sound patterns of the words, which then allowed them to recognize these words when they occurred subsequently in the fluent speech passages.

Since Jusczyk and Aslin's (1995) work, numerous investigations have shown that infants are sensitive to regularities in speech and can use these to segment words from fluent speech. For example, between 6 and 9 months of age, English-learners develop sensitivity to the predominant stress pattern of words in English (Jusczyk, Cutler, & Redanz, 1993). Moreover, they are able to use this information to extract words with the predominant stress pattern from fluent speech (Echols, Crowhurst, & Childers, 1997; Houston, Jusczyk, Kuijpers, Coolen, & Cutler, 2000; Houston, Santelmann, & Jusczyk, in press; Jusczyk, Houston, & Newsome, 1999). Between 7.5 and 10.5 months of age, infants also begin to exploit phonotactic probabilities (Mattys, Jusczyk, Luce, & Morgan, 1999), allophonic cues (Jusczyk, Hohne, & Bauman, 1999), statistical cues (Saffran, Aslin, & Newport, 1996), and coarticulation cues (E. K. Johnson & Jusczyk, 2001) to identify sequences of sounds in fluent speech that correspond to words. The point here is that well before the end of their 1st year, infants have the capacity to extract the sound patterns of potential words from fluent speech.

To be useful in comprehending utterances, the representations of the sound patterns of words that infants encode must persist in long-term memory and generalize to new instances. Although much progress has been made in understanding infants' word segmentation abilities, little information is available regarding their memory for words that they segment from the speech stream. Do infants preserve information about the sound patterns of seg-

Derek M. Houston, Department of Otolaryngology—Head and Neck Surgery, School of Medicine, Indiana University; Peter W. Jusczyk, Departments of Psychology and Cognitive Science, Johns Hopkins University.

Peter W. Jusczyk passed away on August 23, 2001.

This work was supported by National Institute of Child Health and Human Development Research Grant 15795 and National Institute of Mental Health (NIMH) Senior Scientist Award 01490 to Peter W. Jusczyk; it was also supported by NIMH Predoctoral National Research Scientist Award 12232 to Derek M. Houston.

We thank Jenifer Tager, Ann Marie Jusczyk, Amy Gambon-Dennis, Natasha Scheitlin, and Eileen Crowley for their help in recruiting and testing infants. We are also grateful to Ann Marie Jusczyk, Paul Luce, Elizabeth Johnson, Catherine Echols, Carolyn Rovee-Collier, and Catherine Best for comments on earlier versions of this article.

Correspondence concerning this article should be addressed to Derek M. Houston, Department of Otolaryngology—Head and Neck Surgery, DeVault Otologic Research Laboratory, Indiana University School of Medicine, 699 West Drive, Indianapolis, Indiana 46202. E-mail: dmhousto@indiana.edu

mented words longer than the 5-min duration of a testing session in a word segmentation study? If so, what is the nature of infants' word representations? Do the representations contain only abstract, phonological information about the sound patterns of words, or do they contain much more detailed information about the acoustic and/or articulatory characteristics of the specific instances of the words heard? These questions are relevant to understanding how infants begin to build a vocabulary. The representations of sound patterns of words that fluent speakers store in their lexicons allow them to recognize these words even when they are produced by unfamiliar talkers. Does the same hold true for language learners when they begin storing words and building a vocabulary, or is there some developmental period that infants must pass through before they can successfully generalize across different talkers' productions of the same words?

It is possible that infants encode both phonological and indexical information of words in long-term memory and that after hearing different exemplars of the same words repeated in similar contexts, they eventually learn to focus their attention on linguistically relevant properties to identify words. In the word recognition and phonetic structure acquisition (WRAPSA) model, Jusczyk (1993, 1997) proposed that at the first level of infant speech perception, auditory input is picked up by "auditory analyzers," which provide a description of the spectral and temporal features present in the acoustic signal. Jusczyk also postulated that with exposure to speech, infants develop a weighting scheme that gives prominence to features important for understanding the particular language the infant is acquiring. Once the input is recoded and weighted, infants extract the sound patterns of words from fluent speech, and these are matched and stored as representations in the mental lexicon. A fundamental aspect of this model is that it is an exemplar-based system. Instances of words are stored that contain both linguistic and nonlinguistic information, such as the identity and the sex of the talker. In the present investigation, we began to examine whether infants encode both linguistic and talker-specific information in long-term memory by testing infants' ability to recognize words in fluent speech by the original or by a novel talker 1 day after familiarization with productions of these words. Investigating what information infants store in their long-term memory for the sound patterns of words is a first step in understanding how exposure to words may affect the nature of infants' word representations and their word recognition strategies. Moreover, this study provides additional information relevant for understanding how infants become adept at recognizing words across a variety of talkers.

The WRAPSA model and some other exemplar-based models of the mental lexicon represent a shift from a traditional view of the mental lexicon. A traditional view of the mental lexicon assumes that representations of the sound patterns of words consist of phonological descriptions that are stable across different instances of the same word. Word representations are thought to be free of indexical properties—characteristics that provide listeners with information about a talker's identity, sex, emotional state and so forth—that typically vary across different utterances. Several explanations have been offered about how listeners extract phonological information from speech while eliminating the indexical "noise." One suggestion is that there are invariant acoustic properties specifying particular phonemes and that specialized perceptual detectors ignore extraneous information and identify these invariant properties (Blumstein & Stevens, 1980; Fant, 1960; Stevens, 1972; Stevens & Blumstein, 1981). However, the existence of invariant acoustic properties has repeatedly been called into question (Klatt, 1989; Liberman, 1996). Other approaches appeal to a perceptual normalization process that leaves listeners with phonological information. For example, it has been proposed that indexical information is factored out by the perceptual system's accommodation to the characteristics of talkers' vocal tracts and speaking behaviors (e.g., Gerstman, 1968; Shankweiler, Strange, & Verbrugge, 1977) or by listeners' use of acoustic reference points (e.g., F0 and F3) in the speech signal to estimate phonological properties (Strange, 1989; Syrdal & Gopal, 1986). These two views are not incompatible. Listeners could use a combination of the two—the perceptual system may first estimate phonological properties based on reference points and then gradually make adjustments as it acquires talker-specific information (Creelman, 1957; Nusbaum & Morin, 1992). Some researchers have gone so far as to suggest specialized neural assemblies in the brain that are dedicated to extracting the linguistic information from speech and ignoring the indexical properties (Sussman, 1984, 1986). The various approaches in the aforementioned models all share the assumption that talker-specific information must be discarded to arrive at phonologically pure forms.

The traditional view that word representations are phonologically pure has been called into question by findings suggesting that listeners encode indexical properties in their word representations. Specifically, word identification and recognition are affected by talker-specific information. Listeners are better at recognizing the recurrence of a word in a list when the word is presented by the same talker both times than they are if the talkers are different (Bradlow, Nygaard, & Pisoni, 1999; Craik & Kirsner, 1974; Palmeri, Goldinger, & Pisoni, 1993). Likewise, listeners demonstrate same-talker facilitation when identifying words in white noise (Goldinger, 1996; Nygaard, Sommers, & Pisoni, 1994). Goldinger (1996) explored same-talker facilitation effects on implicit and explicit memory of words over delays of 5 min, 1 day, and 1 week. He found same-talker facilitation effects on the implicit memory task (word identification in noise) at all delay intervals, but facilitation effects on the explicit memory task (word recognition) were only present for the 5-min and 1-day delay intervals. Similarly, Church and Schacter have found evidence for voice repetition priming on implicit but not explicit memory tasks (Church & Schacter, 1994; Schacter & Church, 1992). These results suggest that listeners may form implicit memory traces of talker-specific information.

Findings such as these have led some researchers to propose that listeners encode specific, acoustically detailed instances of words in their lexicons rather than abstract, canonical versions of words (Goldinger, 1996, 1998; K. Johnson, 1997; Jusczyk, 1993, 1997). Indeed, Jusczyk's (1993, 1997) WRAPSA model assumes that infants store specific exemplars of the sound patterns of words that contain some acoustic details such as the unique characteristics of individual talkers. Similarly, Remez, Fellowes, and Rubin (1997) have argued that the phonetic properties of words contain talker-specific as well as linguistic information. For example, when presented with sinewave speech, which lacks fundamental frequency information usually associated with voice quality, listeners could nevertheless identify the talker who produced the sentences (Remez et al., 1997). Remez et al. interpreted their findings as indicative of a common phonetic code that listeners use to extract both linguistic and talker-specific information.

Despite growing interest in the nature of word representations in adults, only a handful of investigations have addressed these issues with respect to language learners. Pioneering investigations by Kuhl (1979, 1983) showed that 6-month-olds discriminate vowel contrasts across different talkers, suggesting that at least to some degree, infants can focus on phonetic differences and not react to indexical ones. Likewise, Jusczyk, Pisoni, and Mullennix (1992) found that 2-month-olds could detect syllable changes when these were produced by different talkers. More recently, Houston and Jusczyk (2000) explored the effects of changing talkers on infants' ability to recognize the sound patterns of words. They familiarized 7.5- and 10.5-month-olds with words produced by one talker and then presented the infants with passages produced by a second talker. When the two talkers were of the same sex, 7.5-month-olds showed recognition of the familiarized words. However, when the talkers were of the opposite sex, 7.5-month-olds did not orient longer to the passages, but 10.5-month-olds did. The latter finding suggests that infants' representations of the sound patterns of words become more generalizable toward the end of the first year. Acoustic analyses of the stimuli revealed that the same-sex talkers were more similar to each other with respect to pitch measurements of the target words. Houston and Jusczyk's findings suggest that similarity between talkers plays a role in infants' ability to recognize words across different talkers.

Researchers have only recently begun to investigate infants' long-term memory for the sound patterns of words. Jusczyk and Hohne (1997) found evidence that 8.5-month-olds' memory for the sound patterns of words may persist for as long as 2 weeks. Jusczyk and Hohne familiarized infants with stories produced by different talkers once a day for 10 days during a 2-week period. Two weeks after they had last heard the stories, infants were presented with lists of words that either did or did not occur in the stories. Infants oriented significantly longer to the lists of words from the stories, suggesting that they remembered the words. In a related study, Jusczyk, Hohne, Jusczyk, and Redanz (1993) tested whether another group of infants, familiarized with the same stories but always hearing them produced by a single talker, retained information about the talker. Two weeks after they last heard the stories, infants were once again tested on lists of words that had occurred in the stories, but half of the lists were produced by the familiar talker and the other half by an unfamiliar talker. The infants listened significantly longer to the lists produced by the familiar talker, suggesting that they retained information about that talker's voice characteristics. These findings raise the possibility that infants do encode talker-specific characteristics in their representations of words. However, because the paradigm evaluates how infants respond to an entire list of items and not to individual words, it is difficult to say much on this basis about the nature of the representations of particular words.

The present investigation directly assessed infants' memory for the sound patterns of words with respect to talker-specific information. To our knowledge, this is the first study to investigate infants' long-term memory for both linguistic and indexical properties of spoken words. This approach will provide new information about the nature of infants' word representations following a significant delay. Our approach was twofold. First, we examined the possibility that infants are able to encode the sound patterns of words into long-term memory with a much shorter familiarization period than used in previous investigations (Jusczyk & Hohne, 1997). We used a brief familiarization period and tested infants after a 1-day delay. Second, we explored infants' long-term memory for talker-specific properties of words. Will infants recognize words in fluent speech passages when these are produced by a novel talker 1 day after the infants were familiarized with the words? If so, this might imply that infants encode abstract representations of the phonological characteristics of words, which do not include talker-specific information. Alternatively, a finding that infants do not readily generalize to productions of the familiarized words by a novel talker would suggest that infants' representations of words include indexical information, such as talker-voice characteristics.

To address these issues, three experiments were conducted using a modified version of the Headturn Preference Procedure (Kemler Nelson et al., 1995). Infants were presented with isolated versions of two monosyllabic words in a 3-to-5-min familiarization period, during which each word was heard between 20 and 40 times. On the following day, infants were presented with four passages, two containing the familiarized words and two containing unfamiliar targets. Experiment 1 assessed whether infants retained any information about the familiarized words after a 1-day delay. For this reason, only recordings from a single talker were used. Subsequent experiments involved familiarizing infants with words from one talker and then testing them with passages produced by either another talker (Experiment 2) or by both the novel and the original talker (Experiment 3). These manipulations allowed us to test infants' long-term memory for the sound patterns of words and voices and to determine the effect of talker-specific information on infants' ability to access the sound patterns of words from long-term memory.

Experiment 1

To test infants' long-term memory for the sound patterns of words, we used stimuli identical to and a method similar to those used by Jusczyk and Aslin (1995). They found that 7.5-month-olds but not 6-month-olds oriented significantly longer to passages containing familiarized words than to passages containing unfamiliar target words, suggesting that the older infants can segment and recognize familiarized words in fluent speech contexts. The familiarization period was brief, consisting of between 20 and 40 repetitions of each word, and the test phase followed immediately. Hence, Jusczyk and Aslin did not necessarily test infants' long-term memory for spoken words. In the present experiment, we inserted a 1-day delay between the familiarization and test phases. Otherwise, the experiment was identical to Jusczyk and Aslin's. In contrast to the study conducted by Jusczyk and Hohne (1997), in which infants heard the test words 13 times a day on average for 10 days, the familiarization period in the present study was much less extensive (an average of 30 repetitions of each target). Thus, the present study may prove useful in gauging how little exposure is sufficient for infants to encode the sound patterns of words in long-term memory.

Method

Subjects. Twenty-four American 7.5-month-olds (14 female, 10 male) from monolingual English-speaking families were tested. The infants had

Table 1
Passages Recorded as Stimuli in Test Phases

Target word	Passage
bike	His bike had big black wheels. The girl rode her big bike. Her bike could go very fast. The bell on the bike was really loud. The boy had a new red bike. Your bike always stays in the garage.
dog	The dog ran around the yard. The mailman called to the big dog. He patted his dog on the head. The happy red dog was very friendly. Her dog barked only at squirrels. The neighborhood kids played with your dog.
cup	The cup was bright and shiny. A clown drank from the red cup. The other one picked up the big cup. His cup was filled with milk. Meg put her cup back on the table. Some milk from your cup spilled on the rug.
feet	The feet were all different sizes. This girl has very big feet. Even the toes on her feet are large. The shoes gave the man red feet. His feet get sore from standing all day. The doctor wants your feet to be clean.

a mean age of 33 weeks, 2 days (range = 30 weeks, 6 days to 35 weeks, 4 days; SD = 10 days). Eight additional infants were tested but not included because of failure to complete the full set of test trials due to restlessness and/or crying (4), failure to look for an average of at least 3 s to each stimulus type (2), ear infection (1), and equipment failure (1). All infants were recruited from families living in the greater Buffalo, New York, area.[1]

Stimuli. The stimuli consisted of repetitions of isolated words and passages produced by a single female talker. The words and passages were the same as those used by Jusczyk and Aslin (1995). A passage of six sentences was constructed for each of four words (*cup, dog, feet,* and *bike*). The target word occurred once in each sentence, in variable sentence positions (see Table 1). Within each passage, the target word was always preceded by a different word in each sentence. Across the passages, the words preceding the targets were the same (i.e., *his, red, the, big, old,* and *your*). The words following the targets differed for each target type. We used the same recordings as Jusczyk and Aslin. The female talker had been instructed to read the four passages as if she were speaking to a young child. Next, she was asked to repeat each of the four words (*cup, dog, feet,* and *bike*) successively 15 times with varied intonation, as if speaking to a young infant.

The mean duration of the passages was 19.72 s (ranging from 18.51 s for the *bike* passage to 20.60 s for the *feet* passage; SD = 1 s). The mean duration of the lists was 26.53 s (ranging from 25.84 s for the *feet* list to 27.13 s for the *dog* list; SD = 0.56 s).

Apparatus. The experiment was conducted in a three-sided test booth constructed of pegboard, with 4- × 6-ft (1.2- × 1.8-m) panels on three sides and an open back. This made it possible for an observer to look through one of the existing holes to monitor the infant's headturns. Except for a small section for viewing the infant, the remainder of the pegboard was backed with white cardboard to guard against the possibility that the infant might respond to movements behind the panel. The test booth had a red light and a loudspeaker mounted at eye level on each of the side panels and a green light mounted on the center panel. A white curtain suspended around the top of the booth shielded the infant's view of the rest of the room. A Macintosh Centris 650 computer and response box were located behind the center panel, out of view of the infant. The response box, which was connected to the computer, was equipped with a series of buttons that started and stopped the flashing center and side lights, recorded the direction and duration of headturns, and terminated a trial when the infant looked away for more than 2 s. Information about the direction and duration of headturns and the total trial duration were stored in a data file on the computer.

Procedure. Infants were tested using a version of the Headturn Preference Procedure (Kemler Nelson et al., 1995). On the 1st day, half of the infants were familiarized with one pair of words (*cup* and *dog*), and the other half were familiarized with the other pair (*bike* and *feet*). On the 2nd day, all of the infants were tested on the four passages produced by the original talker. Each infant sat on the lap of a caregiver who was seated on a chair in the center of the test booth. At the beginning of each trial, the center light flashed until the infant oriented to the center. Then, the center light was turned off, and one of the sidelights began flashing. When the infant oriented at least 30° in the direction of the light, the speech stimulus was presented to the same side as the flashing light. The side of presentation was randomized within and across infants and was unrelated to the stimulus conditions. The stimulus continued until the infant looked away for 2 s or until the end of the trial. The amount of time that the infant oriented to the stimulus side while the stimulus was playing was recorded for each trial.

During the familiarization phase, on a given trial each infant was presented with different tokens of one of the two familiarization words. The familiarization words were presented on alternating trials until at least 30 s of looking time was accumulated for each word. The caregivers returned on the following day with their infants. During the test phase, all four passages were presented once in each of four blocks. The order of the passages within each block was randomized. For each subject, two of the passages contained the target words presented during familiarization, and the other two passages contained target words not heard during familiarization. An average orientation time difference between the passages with the familiarized and unfamiliarized target words was taken as an indication that the infants differentiated the two types of passages, presumably because they recognized the familiarized words in the passages.

An observer hidden behind the center panel looked through a peephole and recorded the direction and duration of the infant's headturns, using the response box. The observer was not informed of which items served as familiarization words for a given infant. The loudness level for the samples was set at 72 ± 2 dB (C) SPL using a Quest (Model 215) sound level meter by an assistant who was not involved in the observations. During the experiment, both the observer and the caregiver listened to music over tight-fitting closed headphones (Sony MDR-V600), so they were unaware of which particular stimulus was presented at any given time.

Reliability. The looking times of the infants were computed online by the observer. In order to assess the reliability of our online measures, a second person recoded the looking responses of several of the infants. For

[1] During the years in which data were collected on Experiment 1 and Experiment 2a (1996–1997), the Buffalo metropolitan area population was 85% White and 15% minority. Racial–ethnic and socioeconomic status information was not collected from the study subjects, but subjects were recruited from one suburban hospital and one hospital located in the inner city (in order to maximize minority representation in the samples).

each experiment, the videotape recordings of six subjects were randomly selected for recoding, and correlations between the two codings were calculated: .99 for Experiment 1, .98 for Experiment 2a, .95 for Experiment 2b, and .97 for Experiment 3. These correlations are very high, suggesting that the measures of looking time and direction taken by the original observer were accurate and reliable.

Results and Discussion

The mean orientation times to the passages containing the familiarized words and to the control passages were computed for each infant.[2] The 7.5-month-olds oriented, on average, 9.73 s (SD = 3.68 s) to the familiarized word passages and 8.10 s (SD = 3.67 s) to the control passages (see Figure 1). The difference of the means was 1.63 s (95% CI: $0.72 < 1.63 < 2.54$), and a paired t test revealed that this difference was statistically significant, $t(23) = 3.69$, $p < .01$. Overall, 20 of the 24 infants oriented longer to the passages containing the familiarized words.

The results replicate those reported by Jusczyk and Aslin (1995). In both investigations, 7.5-month-olds oriented significantly longer to the passages containing familiarized words, suggesting that they recognized the words even after a 1-day interval between the familiarization and test phases. The present findings are consistent with those of Jusczyk and Hohne (1997) in that they show that infants retained the sound patterns of words in memory. Moreover, these findings extend the previous work by demonstrating that 7.5-month-olds encoded the sound patterns of words into long-term memory after a relatively short familiarization period. Thus, the results establish that the word segmentation paradigm can be used to explore infants' long-term memory of the sound patterns of familiarized words.

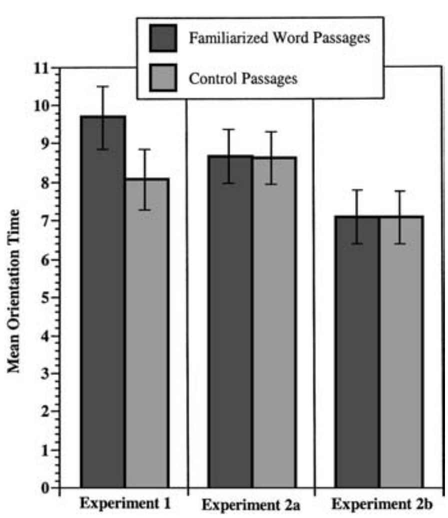

Figure 1. Infants' mean orientation times (in seconds) to test passages 1 day after familiarization in Experiments 1, 2a, and 2b. In Experiment 1, passages were presented to 7.5-month-olds by the original female talker, as in the familiarization phase. In Experiment 2a, 7.5-month-olds were presented with passages by a novel female talker. Experiment 2b was the same as 2a but was administered to 10.5-month-olds. Error bars represent standard errors.

Although the present findings and those of Jusczyk and Hohne (1997) indicate that infants encode speech information into long-term memory, neither study revealed the exact nature of infants' word representations. On the one hand, infants may have encoded both indexical and linguistically relevant properties of the target words. On the other hand, infants could have responded appropriately to the test passages even if their representations only included the linguistically relevant, and not the indexical, information associated with the familiarization words. The possibility that infants encode indexical—namely, talker-specific—properties of speech into their long-term memories for the sound patterns of words was explored in Experiment 2.

Experiment 2a

According to a traditional view of speech perception, representations of the sound patterns of words contain only linguistically relevant information (Halle, 1985; Liberman & Mattingly, 1985). Recently, Houston and Jusczyk (2000) found that 7.5-month-olds could recognize words across some but not all talkers, suggesting that talker-specific information plays a role in an immediate word recognition task. However, it is possible that the biological mechanisms underlying language encoding allow indexical information to decay from memory, leaving only linguistically relevant information. If this is the case, 7.5-month-olds may be better able to generalize words across different talkers when a 1-day delay is imposed between familiarization and test phases.

The traditional–biological view contrasts with views that assume that both mature and developing lexicons are exemplar-based systems (e.g., Goldinger, 1996, 1998; K. Johnson, 1997; Jusczyk, 1993, 1997). Recent findings by Goldinger (1996) have shown that up to 1 week after initial exposure, talker-specific information still facilitates adults' word identification in noise, suggesting that talker-specific information persists in long-term memory. According to the exemplar view, talker-specific information is encoded into word representations and is just as likely as linguistically relevant information to persist in long-term memory. The exemplar view predicts that infants familiarized with words on one day may recognize them when they are produced by a novel talker on the next day only if the instances of the words are

[2] Orientation times tend to decrease across blocks as infants grow less interested in a task. However, decrease in orientation times typically does not interact with the word familiarity effect (e.g., Jusczyk & Aslin, 1995; Jusczyk, Houston, & Newsome, 1999). To verify that the decrease in orientation times across test block did not interact with other variables in the present experiments, an analysis of variance (ANOVA) was performed, combining the orientation time data over all three experiments. In this analysis, the orientation times to the familiar and control passages were calculated for each test block separately. As found in previous studies (Jusczyk & Aslin, 1995; Jusczyk, Houston, & Newsome, 1999), there was a main effect of test block across experiments, $F(3, 300) = 63.06$, $p < .01$, but no significant Test Block × Experiment, $F(9, 300) = 1.41$, $p > .18$, or Test Block × Word Familiarity, $F(3, 300) < 1$, interactions. Also, the Test Block × Experiment × Word Familiarity interaction did not approach significance, $F(9, 300) < 1$. These analyses confirm that although infants tended to decrease their looking times over test blocks, this decrease in looking times did not significantly interact with the magnitude of the difference in orientation times to the familiar versus the control passages in any of the experiments.

sufficiently similar to those from the familiarization phase with respect to both talker-specific and phonological properties.

To test whether talker-specific information is encoded in infants' representations of the sound patterns of words, we familiarized 7.5-month-olds with words (*cup* and *dog* or *bike* and *feet*) produced by one talker on the 1st day, and then we presented them the next day with passages produced by a novel talker. If 7.5-month-old infants' word representations are the same after a 1-day delay as they are immediately after encoding words, then we would expect the same pattern of findings as in Houston and Jusczyk (2000). The infants should be able to recognize words across relatively similar talkers but not across relatively dissimilar talkers. If, after a 1-day delay, indexical properties decay from memory, leaving more abstract word representations, then 7.5-month-olds may demonstrate word recognition across both relatively similar and relatively dissimilar talkers. However, it is possible that talker-specific information plays an important role in infants' ability to access the sound patterns of words from long-term memory. In that case, infants may not exhibit word recognition across different talkers after a 1-day delay, even if the talkers are relatively similar.

In this experiment, we tested 7.5-month-olds' word recognition across two relatively similar female talkers over a 1-day delay. The stimuli were identical to those used in Experiment 1 of Houston and Jusczyk (2000), who found that 7.5-month-olds were able to generalize across the productions of words by the same two female talkers. If 7.5-month-olds demonstrated word recognition in this case, it would support both the possibility that infants' word representations are the same after a 1-day delay and the possibility that their word representations become more abstract. By contrast, a null finding would be consistent with the possibility that infant word recognition depends more on talker-specific properties after a 1-day delay than it does with immediate testing.

Method

Subjects. Twenty-four American 7.5-month-olds (10 female, 14 male) from monolingual English-speaking families were tested. The infants had a mean age of 32 weeks, 6 days (range = 31 weeks, 1 day to 35 weeks, 1 day; SD = 7 days). Two additional infants were tested but not included because of failure to complete the full set of test trials due to crying (1) and experimenter error (1). All infants were recruited from families living in the greater Buffalo, New York, area.

Stimuli. The stimuli consisted of isolated words and passages produced by two different female talkers. The stimuli from Experiment 1, produced by Talker 1, were reused for the present experiment. In addition, a new female talker (Talker 2) recorded the same isolated words (*cup*, *dog*, *feet*, and *bike*) and passages. The same recording conditions and instructions were used for both talkers. For Talker 2, the average duration of the passages was 19.36 s (ranging from 18.96 s for the *cup* passage to 20.12 s for the *feet* passage; SD = 0.52 s). The average duration of the lists was 18.55 s (ranging from 18.34 s for the *feet* list to 18.88 s for the *bike* list; SD = 0.24 s).

Apparatus. This was the same as in Experiment 1.

Procedure. As in Experiment 1, all infants were familiarized with isolated words on one day and then tested with passages the following day. Half of the infants were familiarized with pairs of words produced by Talker 1 and tested on passages produced by Talker 2. The other half were familiarized with pairs of words produced by Talker 2 and tested on passages produced by Talker 1. The isolated words were presented until infants accumulated 30 s of orientation time to each of the two words. For each familiarization talker, half of the infants were familiarized with *cup* and *dog,* and half were familiarized with *bike* and *feet.* All infants were tested on all four passages.

Results and Discussion

The average orientation times to the passages containing the familiarized words and to the control passages were computed for each infant. Infants' mean orientation times were 8.67 s (SD = 3.20 s) to the familiar passages and 8.73 s (SD = 2.97 s) to the control passages (see Figure 1). Only 13 of the 24 infants oriented longer to the familiarized word passages than to the control passages. The difference of the means did not approach statistical significance, $t(23) < 1$ (95% CI: $-1.06 < -0.06 < 0.94$).

The results suggest that infants did not recognize the familiarized words in the passages when these were produced by a novel female talker. These findings stand in contrast to those of Houston and Jusczyk (2000), who used the same stimuli but without 1-day delay. The findings also contrast with the results of Experiment 1, which also imposed a 1-day delay between the familiarization and test phases but used productions from a single talker. To verify that infants in Experiment 2a had behaved differently than those in Experiment 1, we conducted an omnibus ANOVA on the data from both experiments. There was a main effect of familiarity, $F(1, 46) = 6.22$, $p < .02$, indicating that the passages with the familiar words had longer listening times overall, but the main effect of experiment was not significant, $F(1, 46) < 1$. Most important, the interaction of Familiarity $\times$ Experiment was significant, $F(1, 46) = 5.82$, $p < .02$, confirming that the infants in the two experiments had responded differently to the test passages. Thus, despite the fact that Experiment 2a combined conditions from two other experiments in which 7.5-month-olds did display recognition of familiarized words, there was no indication that the infants recognized the familiarized words when these were produced by a novel talker after a 1-day delay.

One possible interpretation of this pattern is that changing talkers and imposing a 1-day delay between the familiarization and the test phases increased the difficulty of the task. The 7.5-month-olds might have recognized the sound patterns of words given either one of these extra demands, but the combination of the two exceeded their processing abilities at this point in development. Following this line of reasoning, older infants with greater cognitive skills and capacity may be able to recognize familiarized words given both a 1-day delay and productions from different talkers. Some support for this view comes from the results of Houston and Jusczyk (2000). They found that 10.5-month-olds but not 7.5-month-olds recognized familiarized words produced by different talkers of the opposite sex. These findings suggest that 10.5-month-olds are more skilled at recognizing words across different talkers than are 7.5-month-olds. Other investigations have revealed that the word segmentation abilities of 10.5-month-olds are more robust than those of 7.5-month-olds (Jusczyk, Hohne, & Bauman, 1999; Jusczyk, Houston, & Newsome, 1999). Hence, 10.5-month-olds will have more capacity to deal with recognizing words across two female talkers and a 1-day delay at the same time. In Experiment 2b, we tested 10.5-month-olds in the same experiment and with the same materials used in Experiment 2a.

Experiment 2b

Method

Subjects. Twenty-four American 10.5-month-olds (12 female, 12 male) from monolingual English-speaking families were tested. The infants had not participated in either Experiment 1 or Experiment 2a. They had a mean age of 46 weeks, 2 days (range = 44 weeks, 5 days to 49 weeks, 0 days; *SD* = 9 days). Nine additional infants were tested but not included because of failure to complete the full set of test trials due to restlessness and/or crying (4), failure to look for an average of at least 3 s to each stimulus type (3), and experimenter error (2). All infants were recruited from families living in the greater Baltimore, Maryland, area.[3]

Stimuli, apparatus, and procedure. These were the same as in Experiment 2a.

Results and Discussion

The average orientation times to the passages containing the familiarized words and to the control passages were computed for each infant. Infants' average orientation times were 7.11 s ($SD =$ 3.25 s) to the familiar passages and 7.11 s ($SD = 3.10$ s) to the control passages (see Figure 1). Fifteen of the 24 infants oriented longer to the familiarized word passages than to the control passages. The means were nearly identical, providing no indication that 10.5-month-olds recognized the familiarized words in the passages, $t(23) < 1$ (95% CI: $-0.63 < 0 < 0.64$).

Like the 7.5-month-olds in Experiment 2a, 10.5-month-olds showed no evidence of recognizing words from different talkers when a 1-day delay was imposed between the familiarization and the test phases. These findings are surprising under the assumption that the problem for 7.5-month-olds in Experiment 2a was simply the additive difficulty of recognizing words across different talkers and over a 1-day delay. If that were true, then the greater word segmentation abilities of 10.5-month-olds (Houston & Jusczyk, 2000; Jusczyk, Hohne, & Bauman, 1999; Jusczyk, Houston, & Newsome, 1999) should have ensured a successful outcome in the present experiment. Instead, it appears that infants' word recognition skills become more dependent on talker-specific properties after a 1-day delay.

How might talker-specific information affect infants' long-term memory for the sound patterns of words? One possibility is that talker-specific information serves to facilitate the access of words from long-term memory. In cases in which no delay occurs between familiarization and test phases, infants' memory for the familiarization words might still be active in short-term memory. When presented immediately after familiarization, instances of the same word produced by a novel talker can be compared to those words in short-term memory. However, after a 1-day delay, indexical information may serve to prime words that are no longer active in short-term memory. Thus, sentences presented on the following day may prime the familiarized words stored in long-term memory when they have the same talker-specific information (i.e., Experiment 1) but not when the talkers are different (i.e., Experiments 2a and 2b). This view is consistent with the view that listeners form implicit representations of acoustic details, which serve to access phonological information (Church & Schacter, 1994). We explored these possibilities in Experiment 3.

Experiment 3

The findings from Experiments 2a and 2b suggest that infants' access to word representations is more dependent on talker-specific information after a 1-day delay than it is immediately after infants have been presented with the words. The purpose of this experiment was to further explore the role of talker-specific information in infant word recognition after a delay. One possibility is that if infants store talker-specific information in long-term memory, then priming them with talker-specific information might improve their retrieval of familiarized words. In situations similar to that of Experiment 1, talker-specific information may help infants access words from long-term memory, provided that they encode such information and maintain it over a 1-day delay. In Experiment 3, we tested the possibility that infants maintain talker-specific properties of words over a 1-day delay by presenting 7.5-month-olds,[4] with two words produced by a female talker on one day and then passages on the next day—some produced by the original talker and others produced by a novel female talker. If infants' memory for talker-specific information persisted over a 1-day delay, then we should have observed differences in orientation times between the passages produced by the two talkers.

As in the previous experiments, two of the passages contained the familiarized words (one passage produced by the original talker and one by the novel talker), and the other two passages contained the unfamiliarized target words (one by the original talker and one by the novel talker). This design allowed us to test the possibility that talker-specific information plays an important role in priming word representations. If talker-specific information activates the memory of the sound patterns of words produced by a particular talker, then the presence of the familiar talker during the test phase may prime both of the familiarized words presented the day before. This priming may enable infants to recognize the familiarized words in both the passages produced by original talker and those produced by the relatively similar novel talker.

Method

Subjects. Thirty-two American 7.5-month-olds (14 female, 18 male) from monolingual English-speaking families were tested. The infants had a mean age of 32 weeks, 1 day (range = 30 weeks, 2 days to 35 weeks, 5 days; *SD* = 8 days). All infants were recruited from families living in the greater Baltimore, Maryland, area.

Stimuli and apparatus. These were the same as in Experiments 2a and 2b.

Procedure. The familiarization phase was identical to that in the previous experiments: Infants were exposed to two words (*cup* and *dog* or *bike* and *feet*) produced by one talker until they accumulated 30 s of orientation time to each word. During the test phase, infants were presented with the same four passages, with each passage occurring once in each of the four randomized test blocks. One change in the present experiment was that two passages were recorded by the female from the familiarization phase

[3] Racial–ethnic data for the sample of the pool of participants in the Infant Language Lab at Johns Hopkins University in Baltimore are as follows: Asian = 0.5%, Black = 21.0%, White = 75.4%, and unknown = 3.1%.

[4] This experiment was conducted with 7.5-month-olds rather than 10.5-month-olds so that the results could be compared with the findings of the previous two experiments. Only 7.5-month-olds were tested in Experiment 1.

(original talker) and the other two passages were recorded by another female (novel talker). There were four passage types: one produced by the original talker and containing the familiarized words (OF), one produced by the novel talker and containing the familiarized words (NF), one control passage produced by the original talker (OC), and one control passage produced by the novel talker (NC). Half of the infants were familiarized with *cup* and *dog*, and half were familiarized with *bike* and *feet*. Within each of these groups, half were familiarized with the words by Talker 1 and the other half were familiarized with the words by Talker 2. Table 2 delineates the eight conditions used to achieve counterbalancing of familiarization words and talkers.

Results and Discussion

The average orientation times to the four passages were computed for each infant. The orientation times were 10.14 s (SD = 3.45 s) to the OF passages, 8.17 s (SD = 2.82 s) to the NF passages, 9.03 s (SD = 3.18 s) to the OC passages, and 6.97 s (SD = 2.64) to the NC passages (see Figure 2). The orientation times were subjected to a repeated-measures ANOVA with word familiarity and talker familiarity as the two variables. There were main effects of both word familiarity, $F(1, 31) = 7.80$, $p < .009$, and talker familiarity, $F(1, 31) = 21.63$, $p < .0001$. There was no interaction between the two factors ($F < 1$), suggesting that infants were able to recognize the familiarized words both in passages produced by the original talker and in the passages produced by the novel talker. In order to more precisely determine whether infants did demonstrate a significant familiarity effect in both talker conditions independently, planned comparisons were used to assess word familiarity for the original-talker condition and the novel-talker condition separately. In agreement with the lack of a Talker Familiarity × Word Familiarity interaction in the ANOVA, the planned comparisons revealed that infants oriented significantly longer to the familiarized word passages than to the control passages in both the original-talker condition, $F(1, 31) = 4.37$, $p < .05$ (95% CI: $-0.05 < 1.11 < 2.26$), and the novel-talker condition, $F(1, 31) = 5.16$, $p < .04$ (95% CI: $0.08 < 1.20 < 2.32$).

The results, first of all, suggest that 7.5-month-olds' memory for talker-specific speech information persists in long-term memory. Infants orient longer to passages when they are presented in the same voice as they were in the familiarization phase than when the voice is different. Second, the findings suggest that indexical information plays a role in accessing words from long-term memory. In contrast to Experiment 2, infants showed word recognition across different talkers and over a 1-day delay, suggesting that the presence of the voice from the familiarization phase primed infants' long-term memory for the familiarization words. Once the familiarized words were primed, infants were able to recognize them not only in the passages produced by the familiar talker but also in the passages produced by another talker of the same sex.

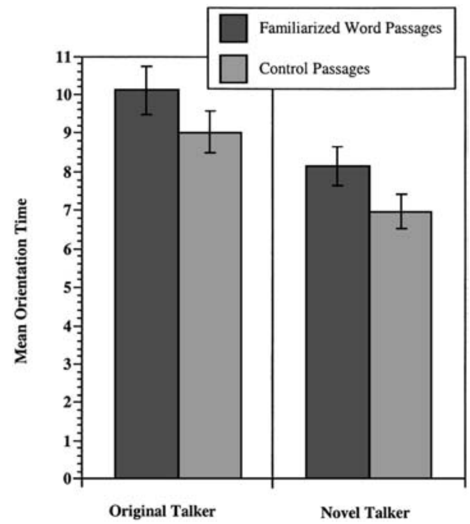

Figure 2. Displays of 7.5-month-olds' mean orientation times (in seconds) to test passages in Experiment 3. Passages were presented by both the original female talker from the familiarization phase and a novel female talker. There was a 1-day delay between the familiarization and test phases. Error bars represent standard errors.

General Discussion

The findings from the experiments reported here suggest that by 7.5 months of age, infants encode the sound patterns of words into long-term memory. Experiment 1 revealed that 7.5-month-olds

Table 2
Conditions Used to Counterbalance Familiarization Words and Talkers in Experiment 3

Condition	Isolated words (familiarization phase)	Words/passage type (test phase)			
		OF	NF	OC	NC
1	cup, dog	cup	DOG	bike	FEET
2	cup, dog	dog	CUP	feet	BIKE
3	CUP, DOG	DOG	cup	FEET	bike
4	CUP, DOG	CUP	dog	BIKE	feet
5	bike, feet	bike	FEET	cup	DOG
6	bike, feet	feet	BIKE	dog	CUP
7	BIKE, FEET	BIKE	feet	CUP	dog
8	BIKE, FEET	FEET	bike	DOG	cup

Note. Infants were familiarized with isolated words presented by either Talker 1 (lowercase letters) or Talker 2 (uppercase letters). Infants' word recognition was tested by presenting four passage types: (a) passages containing familiarized words produced by the original talker (OF), (b) passages containing familiarized words produced by the novel talker (NF), (c) control passages produced by the original talker (OC), and (d) control passages produced by the novel talker (NC).

recognized words that they heard spoken by the original talker from the previous day. However, Experiments 2a and 2b indicated that infants did not display similar recognition of familiarized words on the next day when these words were produced by a novel talker of the same sex. These results contrast with those of Houston and Jusczyk (2000), who found that 7.5-month-olds generalized across productions of the same words by these talkers when no delay intervened between familiarization and test phases. The pattern of findings suggests that talker-specific information appears to be more important after a 1-day delay than it is immediately following familiarization. In Experiment 3, 7.5-month-olds showed an overall preference for passages produced by the talker heard during familiarization on the previous day. This finding suggests that infants retain information about talker voice characteristics in long-term memory, confirming the findings reported by Jusczyk et al. (1993) for infants who had had 10 days of exposure to a particular voice. Furthermore, unlike in Experiment 2, infants demonstrated recognition of familiarized words in passages produced by an unfamiliar talker as well as in those produced by the familiar talker. The pattern of findings suggest that hearing the voice of the familiar talker during the test phase helped infants to recall the familiarized words from memory even when these were produced by an unfamiliar talker.

The present findings support the notion that linguistic and nonlinguistic aspects of speech are integrally related. Investigations of adult speech perception suggest that talker variability affects listeners on a range of tasks, such as phoneme identification (Green, Tomiak, & Kuhl, 1997; Mullennix & Pisoni, 1990), serial recall of words (Goldinger, Pisoni, & Logan, 1991; Martin, Mullennix, Pisoni, & Summers, 1989), word recognition (Craik & Kirsner, 1974; Palmeri et al., 1993), and word identification (Goldinger, 1996; Nygaard et al., 1994). Similarly, our results show that talker-specific information does have a bearing on whether infants are likely to recognize particular words after a 1-day delay. Moreover, the fact that infants in Experiment 3 listened longer to the familiar voice that they had heard on the previous day than to a novel voice indicates that they did retain information about talker voice characteristics. Might infants encode any type of co-occurring indexical information from their environment into their word representations? Investigations manipulating visual or other contexts between familiarization and test phases would be necessary to answer this question. In one investigation, Sheffert and Fowler (1995) presented adult listeners with videotapes of talkers uttering single words. The investigators varied both visual and auditory information and found that changes in talker voice had a significant effect on word identification, whereas changes made to visual properties, such as talker's face, had only a very weak effect on word identification. Listeners also appear to separate important speech information from other auditory information. For example, changes in amplitude between familiarization and test do not affect perceptual identification of words (Sommers, Nygaard, & Pisoni, 1994).

Our findings fit with the view that infants encode indexical information, such as talker voice characteristics, into their representations of the sound patterns of words. Indeed, the findings are consistent with the possibility that infants' representations of words are exemplar-based (Goldinger, 1996, 1998; Jusczyk, 1993, 1997). Nevertheless, a careful consideration of the present results indicates that they do not provide definitive proof of the latter view, because the type of experimental paradigm used here does not test whether infants show better recognition of an exact token of a word heard during familiarization. Rather, infants showed a greater ability to generalize to new tokens of a word produced by the familiar talker than by an unfamiliar talker. In this sense, the present findings establish that infants' representations of the sound patterns of words include some talker-specific information. However, such representations could be specific to a talker, yet general, in the sense that they are not individual tokens per se, but still summarize the average characteristics of a given word produced by this particular talker. Thus, a true test of the exemplar-based view requires evidence from a different kind of paradigm that allows one to assess whether performance with the exact tokens heard during familiarization exceeds that obtained with other tokens produced by the same talker. In the meantime, it is worth noting that the present data are also compatible with the view that listeners form implicit memory traces of indexical properties of words, which are linked to implicit memory traces of the phonological properties of words (Church & Schacter, 1994).

The present findings are also compatible with views that posit that both phonological and indexical information are represented by a common code. In one such view, Remez et al. (1997) have proposed that listeners extract both types of information via a common phonetic code that specifies allophonic variations of specific utterances. Similarly, an ecological approach to speech perception posits that dynamic articulatory properties afford apprehension of both talker-specific and linguistically relevant information (e.g., Best, 1994; Fowler, 1986). The present findings suggest that infants encode talker-specific information but do not address what form this information takes. Understanding what information infants encode only goes so far in providing a description of infants' representations of the sound patterns of words. Assessing infants' perception of articulatory and acoustic–phonetic information will help to provide a more complete description of the ontology of infants' word representations.

Another issue that merits discussion is the discovery that talker changes are more disruptive to infants' recognition of words after a 1-day delay than with no delay (as in Houston & Jusczyk, 2000). A similar finding has been reported by Hartshorn, Rovee-Collier, and their colleagues for infants' memory of visual displays (Hartshorn et al., 1998). In the standard testing paradigm used in many of these authors' investigations, infants are taught that they can control the movements of a mobile by kicking their legs (2- to 6-month-olds) or the movements of toy trains by pushing levers (6-month-olds and older). This training phase immediately follows a baseline measure in which the infants' kicks and lever pushes do not affect the mobile or toy trains. During a long-term retention test that occurs (depending on the experiment) days, weeks, or longer after the baseline and training phases, the infants are reintroduced to the mobile or trains, and as in the baseline phase, their behavior does not affect the mobile or toy trains. Infants evidence long-term memory for the toys and training session if their response rate (frequency of kicking or lever pushing) during the long-term retention test is significantly greater than it was during the baseline phase. One such line of investigation explored the effects of changing contexts on infant memory over short and long delay intervals. Changing the immediate contextual surroundings (e.g., the rooms where training and testing were conducted) between training and test phases did not impair infants' memory after short delay intervals, but it did affect their memory after long intervals. This pattern of findings held for 3-, 9-, and 12-month-olds (Hartshorn et al., 1998) but not for 6-month-olds (Borovsky & Rovee-Collier, 1990; Shields & Rovee-Collier, 1992). Although the speech studies and these visual memory studies are not entirely analogous—the delays were longer in the visual memory investigations, and changing talkers in the present investigations can only loosely be conceived of as "context effects"—the similar patterns may be revealing of the general role that contextual details play in infants' long-term retrieval of visual and speech information. Evaluating infants' memory for the sound patterns of words across different delay intervals would help in assessing the extent to which memory processes are similar across these two modalities.

How might talker-specific information facilitate infant word recognition after a delay? First, consider what may happen during immediate testing. As suggested earlier, when test passages immediately follow familiarization with isolated words, information about the sound patterns of the familiarized words may still be active in short-term memory. Activation of word representations may help infants to notice the match between these words and their counterparts in the passages, even when the latter are produced by a novel talker. However, as infants begin to encode traces of sound patterns of the words as representations in long-term memory, items that mismatch on many acoustic dimensions (e.g., talker-specific information, speaking rate differences, coarticulatory influences from surrounding words) may be much less activated by the speech of a novel talker producing those same words. Consequently, a slight increase in activation afforded by a match in talker-specific information from the familiar talker could spell the difference in whether the familiarized word is recognized or not after a delay.

Our findings suggest that when beginning to segment words, infants may have difficulty recognizing a familiar word produced by a novel talker. In this respect, such infants' word recognition abilities are not as robust as those of more mature users of the language. However, Houston and Jusczyk (2000) noted some improvement between 7.5 and 10.5 months in infants' ability to generalize immediately from one talker's productions of words to those of another talker of the opposite sex. Certainly, by the time infants have begun to produce many words of their own, they do not appear to have difficulties in recognizing familiar words produced by novel talkers. Is it possible that infants eventually learn to encode only linguistically relevant information in their word representations and discard indexical information such as talker-specific properties? This seems unlikely in view of the fact that adults appear to encode talker-specific information in their word representations (Goldinger, 1996, 1998).

Assuming that infants do retain talker-specific information in their representations of the sound patterns of words, what factors influence their ability to generalize these representations to words produced by novel talkers? Houston (1999) has begun to explore this issue. He identified several factors that seem to promote infants' abilities to generalize to new talkers. One such factor has to do with how the novel voice relates to previous ones with which infants have had experience. In particular, infants are more apt to generalize to a new talker whose voice falls within, rather than outside, the perceptual space defined by the voices of familiar talkers. An additional factor that may facilitate infants' generalization of words to new talkers is common reference. After all, many words that infants hear are used to pick out particular objects in the real world. Even if infants encode talker-specific information into their representations of words, different exemplars of the same word will often be associated with objects that share a great deal of visual similarity in the world. When exemplars that differ chiefly in indexical rather than phonetic properties pick out visually similar objects, learners may be induced to give less attentional weight to indexical properties when listening for meaningful distinctions. Thus, as development proceeds, language users may continue to encode indexical information but learn to focus their attention on the phonetic properties most relevant for language, allowing them to more easily recognize linguistic equivalence (K. Johnson, 1997).

In conclusion, the present experiments indicate that infants encode information about the sound patterns of words into long-term memory, even after a relatively brief exposure period. In addition, the representations of words that infants begin storing in long-term memory appear to include indexical information, relating to talker-voice characteristics. The storage of such talker-specific word representations does, at least initially, impair the ability of infants to generalize their representations of words to those produced by unfamiliar talkers.

References

Best, C. T. (1994). Learning to perceive the sound pattern of English. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in infancy research* (Vol. 9, pp. 217–304). Norwood, NJ: Ablex.

Bloom, L. (1975). *One word at a time.* The Hague, the Netherlands: Mouton de Gruyter.

Blumstein, S. E., & Stevens, K. N. (1980). Perceptual invariance and onset spectra for stop consonants in different vowel environments. *Journal of the Acoustical Society of America, 67,* 648–662.

Borovsky, D., & Rovee-Collier, C. (1990). Contextual constraints on memory retrieval at six months. *Child Development, 61,* 1569–1583.

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

Church, B. A., & Schacter, D. L. (1994). Perceptual specificity of auditory priming: Implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 521–533.

Clark, E. V. (1973). What's in a word? On the child's acquisition of semantics in his first language. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 65–110). New York: Academic Press.

Craik, F. I. M., & Kirsner, K. (1974). The effect of speaker's voice on word recognition. *Quarterly Journal of Experimental Psychology, 26,* 274–284.

Creelman, C. D. (1957). Case of the unknown talker. *Journal of the Acoustical Society of America, 29,* 655.

Echols, C. H., Crowhurst, M. J., & Childers, J. (1997). Perception of rhythmic units in speech by infants and adults. *Journal of Memory and Language, 36,* 202–225.

Fant, C. G. M. (1960). *Acoustic theory of speech production.* The Hague, the Netherlands: Mouton de Gruyter.

Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics, 14,* 3–28.

Gerstman, L. (1968). Classification of self-normalized vowels. *IEEE Transactions on Audio and Electroacoustics, ACC-16,* 78–80.

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.

Goldinger, S. D., Pisoni, D. B., & Logan, J. S. (1991). On the nature of talker variability effects on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 152–162.

Green, K. P., Tomiak, G. R., & Kuhl, P. K. (1997). The encoding of rate and talker information during phonetic perception. *Perception & Psychophysics, 59,* 675–692.

Halle, M. (1985). Speculations about the representation of words in memory. In V. A. Fromkin (Ed.), *Phonetic linguistics: Essays in honor of Peter Ladefoged* (pp. 101–114). San Diego, CA: Academic Press.

Hartshorn, K., Rovee-Collier, C., Gerhardstein, P., Bhatt, R. S., Klein, P. J., Aaron, F., et al. (1998). Developmental changes in the specificity of memory over the first year of life. *Developmental Psychobiology, 33,* 61–78.

Houston, D. M. (1999). *The role of talker variability in infant word representations.* Unpublished doctoral dissertation, Johns Hopkins University.

Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1570–1582.

Houston, D. M., Jusczyk, P. W., Kuijpers, C., Coolen, R., & Cutler, A. (2000). Cross-language word segmentation by 9-month-olds. *Psychonomic Bulletin & Review, 7,* 504–509.

Houston, D. M., Santelmann, L., & Jusczyk, P. W. (in press). English-learning infants' segmentation of trisyllabic words from fluent speech. *Language and Cognitive Processes.*

Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language, 44,* 548–567.

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 145–165). San Diego, CA: Academic Press.

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., Cutler, A., & Redanz, N. (1993). Preference for the predominant stress patterns of English words. *Child Development, 64,* 675–687.

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

Jusczyk, P. W., Hohne, E. A., & Bauman, A. (1999). Infants' sensitivity to allophonic cues for word segmentation. *Perception & Psychophysics, 61,* 1465–1476.

Jusczyk, P. W., Hohne, E. A., Jusczyk, A. M., & Redanz, N. J. (1993). Do infants remember voices? *Journal of the Acoustical Society of America, 93,* 2373.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology, 39,* 159–207.

Jusczyk, P. W., Pisoni, D. B., & Mullennix, J. (1992). Some consequences of stimulus variability on speech processing by 2-month-old infants. *Cognition, 43,* 253–291.

Kemler Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., & Gerken, L. A. (1995). The Headturn Preference Procedure for testing auditory perception. *Infant Behavior & Development, 18,* 111–116.

Klatt, D. H. (1989). Review of selected models of speech perception. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 169–226). Cambridge, MA: MIT Press.

Kuhl, P. K. (1979). Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories. *Journal of the Acoustical Society of America, 66,* 1668–1679.

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

Liberman, A. M. (1996). Some assumptions about speech and how they changed. In A. M. Liberman (Ed.), *Speech: A special code* (pp. 1–44). Cambridge, MA: MIT Press.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition, 21,* 1–36.

Martin, C. S., Mullennix, J. W., Pisoni, D. B., & Summers, W. V. (1989). Effects of talker variability on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 676–684.

Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology, 38,* 465–494.

Mullennix, J. W., & Pisoni, D. B. (1990). Stimulus variability and processing dependencies in speech perception. *Perception & Psychophysics, 47,* 379–390.

Nusbaum, H. C., & Morin, T. M. (1992). Paying attention to differences among talkers. In Y. Tohkura, E. Vatikiotis-Bateson, & Y. Sagisaka (Eds.), *Speech perception, production, and linguistic structure* (pp. 113–134). Tokyo: OHM Publishing.

Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science, 5,* 42–46.

Palmeri, T. J., Goldinger, S. D., & Pisoni, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 309–328.

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.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996, December 13). Statistical learning by 8-month-old infants. *Science, 274,* 1926–1928.

Schacter, D. L., & Church, B. A. (1992). Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 915–930.

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.

Sheffert, S. M., & Fowler, C. A. (1995). The effects of voice and visible speaker change on memory for spoken words. *Journal of Memory and Language, 34,* 665–685.

Shields, P. J., & Rovee-Collier, C. (1992). Long-term memory for context-specific category information at 6 months. *Child Development, 63,* 245–259.

Sommers, M. S., Nygaard, L. C., & Pisoni, D. B. (1994). Stimulus variability and spoken word recognition: I. Effects of variability in speaking rate and overall amplitude. *Journal of the Acoustical Society of America, 96,* 1314–1324.

Stevens, K. N. (1972). The quantal nature of speech. In J. E. E. David & P. B. Denes (Eds.), *Human communication: A unified view* (pp. 51–66). New York: McGraw-Hill.

Stevens, K. N., & Blumstein, S. E. (1981). The search for invariant acoustic correlates for phonetic features. In P. D. Eimas & J. L. Miller (Eds.), *Perspectives on the study of speech* (pp. 1–38). Hillsdale, NJ: Erlbaum.

Strange, W. (1989). Dynamic specification of coarticulated vowels spoken in sentence context. *Journal of the Acoustical Society of America, 85,* 2135–2153.

Sussman, H. M. (1984). A neuronal model for syllable representation. *Brain and Language, 22,* 167–177.

Sussman, H. M. (1986). A neuronal model of vowel normalization and representation. *Brain and Language, 28,* 12–23.

Syrdal, A. K., & Gopal, H. S. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America, 79,* 1086–1100.

van de Weijer, J. (1998). *Language input for word discovery* (Vol. 9). Nijmegen, the Netherlands: Max Planck Institute.

Woodward, A. L., & Markman, E. M. (1997). Early word learning. In W. Damon, D. Kuhn, & R. Siegler (Eds.), *Handbook of child psychology* (5th ed., Vol. 2, pp. 371–420). New York: Wiley.

Woodward, J. Z., & Aslin, R. N. (1990, April). *Segmentation cues in maternal speech to infants.* Paper presented at the 7th biennial meeting of the International Conference on Infant Studies, Montreal, Quebec, Canada.

Received July 16, 2001
Revision received November 25, 2002
Accepted June 2, 2003 ■