

Abstraction and specificity in preschoolers' representations of novel spoken words

Cynthia Fisher

Caroline Hunt

Kyle Chambers

University of Illinois at Urbana-Champaign

and

Barbara Church

SUNY at Buffalo

Running head: Preschoolers' representations of novel words

Corresponding author: Cynthia Fisher, Department of Psychology, University of Illinois, Champaign, IL 61820

Acknowledgements: This research was partially supported by NIH Grant 1 R55 HD/OD34715-01, NSF Grant SBR 98-73450, and by the Research Board of the University of Illinois at Urbana-Champaign. We thank the parents and children who participated in these studies, and Madelena McClure, Jennifer Lanter, Azadeh Ghaffari, Mychelle Denny, and all the students in the Language Acquisition Lab who assisted in data collection. We also thank Renée Baillargeon, Susan Garnsey, LouAnn Gerken, and two anonymous reviewers for helpful comments on a previous version of this paper.

## Abstract

Four experiments explored long-term auditory priming for novel words (nonwords) in preschoolers. In Experiment 1, 2.5-year-olds more accurately identified novel words that had been presented just twice in an initial study phase than nonwords that had not been presented, showing auditory priming for nonwords. Experiments 2, 3, and 4 revealed that the sound representations underlying auditory priming in young children, as in adults, include both abstract and token-specific information about the sounds of new words. In Experiment 2, 3-year-olds showed priming for studied nonsense syllables that changed both token and recorded context from study to test, compared to entirely new test syllables. In Experiment 3, 3-year-olds more accurately identified nonsense syllables that were the same tokens in the same context at study and test than syllables that changed token and context from study to test. In Experiment 4, 3-year-olds more accurately identified the same-token syllables from Experiment 3, even when those syllables were presented in isolation, spliced out of their original contexts. Thus children's rapidly formed representations of new spoken words include both components abstract enough to identify the same sound sequence across changes in word token and changes in phonetic context, and components specific to the originally presented token. We argue that the powerful perceptual learning mechanism underlying auditory word priming has the right properties to play a central role in the development of the auditory lexicon.

In order to learn to identify spoken words, children must establish representations in long-term memory of the sound-patterns of words in their language, and use those representations to identify familiar words in connected speech. In the process a number of perceptual problems must be solved, including the identification of word boundaries and compensation for many sources of acoustic variation in speech. What learning mechanisms could permit the development of representations for word identification? Consideration of familiar aspects of the word learning problem yields a number of constraints on the kinds of learning mechanisms that could serve this purpose.

First, children must be able to learn about the sounds of words before attaching a meaning to each sound pattern. On virtually every theory of word learning, the learner collects information about a word's linguistic and extra-linguistic contexts in order to determine its meaning (e.g., Fisher, Hall, Rakowitz, & Gleitman, 1994; Pinker, 1989; Woodward & Markman, 1997). Thus word learning requires a memory system for representing words as sound patterns without fixed meaning.

Second, children must form representations of the sounds of words on very brief exposure. Two-year-olds can sometimes pick up both the sound and at least part of a new word's meaning on a single trial (e.g., Carey & Bartlett, 1978; Heibeck & Markman, 1987). Even 13-month-olds can fast-map a new word after hearing it only a few times (e.g., Woodward, Markman, & Fitzsimmons, 1994). This fast mapping requires (among other things) a learning mechanism that establishes a robust long-term representation of a new word sound with very little exposure.

Third, children's sound representations must allow them to match previously encoded sound patterns with new instances of the same sound patterns in ordinary connected speech. One perceptual problem posed by connected speech is the enormous amount of acoustic variability among tokens of the same word. Factors such as the surrounding phonetic and prosodic context, speaker's voice, and speech rate radically alter the acoustic properties of speech (e.g., Fisher & Tokura, 1996a, 1996b; Klatt, 1980; Lively, Pisoni, & Goldinger, 1994; Miller & Volaitis, 1989;

Mullennix & Pisoni, 1990; Nusbaum & Goodman, 1994). To identify words as such, across contexts and acoustic variants, speech identification processes must readily abstract across all these kinds of variability. This fits our qualitative impression that speech perception is categorical (though see Massaro, 1994 for arguments that this qualitative impression is misleading).

Fourth, there is reason to suspect that context-sensitive and detailed acoustic-phonetic representations of the sounds of words are a necessary part of a system for spoken word identification. That is, we will argue that word recognition processes cannot entirely discard information about speech sound variability across contexts. The present research bears on this question, so we will discuss this requirement at greater length.

The traditional approach to the problem of variability has been to assume that listeners normalize speech, stripping away acoustic variation due to voices and other influences, to reveal an abstract, context-free, sound pattern for each word. While it is clear that listeners must readily abstract over variability in the sounds of words, it is not so obvious that this process of abstraction must involve the loss of more detailed information (e.g., Goldinger, 1998). Languages differ in their phonetic as well as their phonological systems (e.g., Farnetani, 1997; Keating, 1985,1990; Pierrehumbert, 1990). Speech sounds are coarticulated—they are altered to some degree by nearby sounds. Some of this context-dependent variability may be a natural (and therefore universal) result of vocal production constraints. Language-specific variations in these context effects, however (e.g., Farnetani, 1997; Keating, 1985, 1990; Pierrehumbert, 1994), suggest that listeners must learn how speech sounds are affected by various contexts in their language. If so, then quite detailed and context-sensitive representations of spoken words should be needed to acquire the phonological and phonetic systems of the native language.

To illustrate the problem: vowels tend to be lengthened before voiced rather than unvoiced consonants (e.g., Chen, 1970; Crystal & House, 1988), and vowel duration is interpreted by listeners as a cue to the voicing of the following consonant (e.g., Gordon, 1989; Klatt, 1976). This context effect, among many others, may follow in part from natural

tendencies in speech articulation (e.g., Chen, 1970; Maddieson, 1997), and in speech perception (e.g., Kluender, Diehl, & Wright, 1988). But context-dependent duration differences are not uniform across languages. Keating (1985) reports that vowel length differences attributable to consonant voicing are large and systematic in English, but not in Czech or Polish. She argues, based on these and other data, that part of knowing any language is developing a quantitative estimate of how much vowels change their duration in various phonetic contexts.

To make matters even more complex, just as vowel length differences before voiced and unvoiced consonants vary across languages, so do other influences on vowel length: Utterance-final lengthening is much greater in English than in Japanese (e.g., Campbell, 1992; Fisher & Tokura, 1996a), for example, and duration is a major determinant of syllable stress in English but not in Japanese (e.g., Beckman & Pierrehumbert, 1986; Takeda & Ichikawa, 1994). Vowel length itself is also used contrastively in many languages, and the extent to which listeners rely on vowel length for vowel identification differs across languages and even dialects of the same language (e.g., Miller & Grosjean, 1997). Any acoustic quantity, such as segment duration, is interpreted simultaneously as evidence about the identity of the local segment, its phonetic context, and its position in an intonational phrase (e.g., Fisher & Tokura, 1996a; Gordon, 1989; Klatt, 1976). This extended example concerned one type of variation—in vowel length—but the same argument can be made for virtually every way in which speech sounds vary, for consonants as well as for vowels (see, e.g., Keating, 1990; Klatt, 1980; Pierrehumbert, 1994).

Thus, part of learning a language is collecting detailed information about how sound patterns are realized in different contexts. Listeners must be able to abstract over variation to identify tokens of the same sound sequence, but retain enough information about token variation and context to learn how speech sounds are affected by their contexts in a particular language.

There is another reason why children need to encode specific information about the contexts in which each sound pattern occurs. Thus far we have written as if the child already knows the sound pattern of a word, and is trying to identify it in various guises. But an important part of identifying spoken words is locating their boundaries in connected speech.

Spoken words are not routinely set apart by pauses or other acoustic cues to their boundaries. Languages tend to offer probabilistic pre-lexical cues to the location of word boundaries, including stress pattern, phonotactic regularities, and the duration of segments in different word positions (e.g., Gow & Gordon, 1995; Cutler & Norris, 1988; Nakatani & Dukes, 1977; Nakatani & Schaffer, 1978; Quené, 1992). But pre-lexical cues leave considerable ambiguity about the location of word boundaries in connected speech, and most such cues are language-specific (e.g., Cutler & Norris, 1988; Mattys, Jusczyk, Luce, & Morgan, 1999; McQueen, 1998; Nakatani & Dukes, 1977). For these reasons, theories of word recognition assume that adults find word boundaries in the speech stream at least in part by recognizing the sound patterns of familiar words (e.g., Dahan & Brent, 1999; Klatt, 1980; Marslen-Wilson, 1987; McClelland & Elman, 1986; McQueen, Cutler, Briscoe, & Norris, 1995). Recent research suggests that word segmentation begins in infancy in the same way, with a distributional analysis of sound patterns (Goodsitt, Morgan, & Kuhl, 1993; Jusczyk, 1997; Saffran, Aslin, & Newport, 1996). Sequences of sounds that frequently repeat against a variable background become wordlike perceptual units.

To find word boundaries via distributional analysis, the learner must operate without an unbreakable initial commitment to particular units of speech. Information about sound patterns in context must be represented, and used abstractly to identify an old pattern in a new context, while still retaining enough specific information about context to detect the repeating sequences of sounds that make up words.

Moreover, distributional patterns in speech must be detected at multiple levels of analysis. Languages set limits on which sounds can occur in sequence; these are known as phonotactic regularities. There are widespread similarities across languages in these regularities, supporting proposed universal principles of syllabic organization such as sonority sequencing (e.g., Harris, 1994). But many aspects of phonotactics are language-specific. In English, for example, words can begin with the consonant sequence [kl] but not [tl]; [tl] is a perfectly acceptable onset in the North American language Tlingit. These language-specific phonotactic patterns are part of what listeners learn about their native language.

Speakers and listeners are also sensitive to less absolute regularities than the interdiction on initial [tl] in English. Adults take longer to identify and to produce nonwords whose component sound sequences are of low frequency (e.g., Vitevitch & Luce, 1999). By about 9 months of age infants show sensitivity to the phonotactic regularities and frequencies of their languages (e.g., Mattys et al., 1999; Jusczyk, Luce, & Charles-Luce, 1994). Recent evidence suggests that the adult's language processing system remains sensitive to local phonotactic frequencies established within the context of an experiment (Dell, Reed, Adams, & Meyer, 2000). Thus from infancy to adulthood, listeners pick up frequency-sensitive information about how various units of sound are sequenced in connected speech. Similar distributional evidence is needed to yield both words and sub-lexical patterns. This learning, again, requires a mechanism which both abstracts across contextual variation and encodes specific information about context.

We have argued that known characteristics of the word-learning situation suggest certain requirements for learning mechanisms that could create representations for word identification. Any viable mechanism must (1) build sound representations without requiring prior knowledge of word meaning; (2) establish long term sound representations on one or few exposures; (3) abstract across variability; and (4) retain enough specific contextual and acoustic information to permit learning of distributional patterns and acoustic-phonetic variation in a particular language.

Recent work in the adult memory literature points to a learning mechanism that has just these properties. Research examining long-term auditory word priming in adults has revealed basic implicit memory mechanisms that create and update adults' representations of the sounds of words to reflect auditory experience (e.g., Church & Schacter, 1994; Goldinger, 1996; Schacter & Church, 1992). Each time a word is heard, a lasting representation of its sound pattern is encoded that facilitates later identification of the same word. Fundamentally similar learning principles operate within many perceptual learning domains, including the visual identification of faces (e.g., Althoff & Cohen, 1999) and of written words (e.g., Light, La Voie & Kennison, 1995; Rueckl & Olds, 1993); the nature of the perceptual representations created by

these mechanisms varies with the domain. Considerable evidence suggests that the perceptual learning that underlies auditory word priming has exactly the properties which we argued above are required to create representations for word identification.

First, long-term auditory priming appears to depend on auditory, not semantic, representations. Several findings show that long-term priming is modality specific (e.g., Jackson & Morton, 1984; McClelland & Pring, 1991), suggesting that the mediating representations are specifically auditory. Long-term auditory priming is also unaffected by whether the study task focuses listeners' attention on the sound or the meanings of the words (e.g., Church & Schacter, 1994; Schacter & Church, 1992). This suggests that merely listening to an item produces the priming, without requiring attention to its meaning. Finally, long-term auditory priming can be found for nonsense words, with no previously established meaning (e.g., Goldinger, 1998; and the current experiments).

Second, repetition priming happens fast (on one trial), and can have very long-lasting effects. Single-trial priming for written word identification can be measured a day, a week, or a year later (e.g., Jacoby & Dallas, 1981; Sloman, Hayman, Ohta, Law & Tulving, 1988), and auditory word priming has been measured after a week's delay (Goldinger, 1996).

Third, auditory word priming spans a variety of acoustic-phonetic changes in the word token from study to test, such as a change in pronunciation, voice, or context. Thus abstract matches occur, encompassing the variability found in natural speech (e.g., Church, 1995; Church, Dell, & Kania, 1996; Church & Schacter, 1994; Goldinger, 1996; Poldrack & Church, 1997; Schacter & Church, 1992; Sheffert, 1998; Sommers, 1999).

Fourth, under many circumstances auditory word priming is reduced by acoustic-phonetic changes from study to test. Priming is reduced when the primed word is spoken in a different voice (Church & Schacter, 1994; Goldinger, 1996, 1998; Pilotti et al., 2000; Schacter & Church, 1992; Sheffert, 1998), or when fundamental frequency (Church & Schacter, 1994; Sommers, 1999), formant frequency, or speaking rate (Church, 1995; Sommers, 1999) changes slightly from study to test. Reductions in priming due to a change in speaker's voice have been

measured a week later (e.g., Goldinger, 1996). Thus the rapid and long-lasting representational changes underlying priming reflect acoustic-phonetic details specific to particular tokens of spoken words.

For all these reasons, auditory word priming in adults suggests the operation of a learning mechanism with the properties necessary to support acquisition of the sound patterns of words. The encoding and retrieval of sound pattern information (1) seems not to depend on access to word meaning, (2) occurs on little exposure and leaves a lasting trace, (3) allows abstract matches across token variability, but at the same time (4) retains token-specific details. On this view, each encounter with spoken language adds perceptual information to the language processing system, permitting it to adapt quickly to recent experience while continuing to reflect accumulated experience. This continual addition of new perceptual information to the memory representations for speech recognition permits us to adapt to new words, voices, accents, or dialects (see, e.g., Church & Schacter, 1994; Goldinger, 1998; Nusbaum & Goodman, 1994; Pisoni, 1992). The central hypothesis of this paper is that the same mechanism could create those representations in childhood.

Preliminary studies exploring this continuity of learning hypothesis (Church & Fisher, 1998) revealed that 2-, 2.5-, and 3-year-olds showed patterns of long-term auditory priming very similar to those found in adults. In an elicited imitation task, children more accurately identified and repeated mildly low-pass filtered words that they had heard presented once a few minutes before, than words that were not played previously. Given similar word identification baselines accomplished through differential low-pass filtering of familiar words, the magnitude of auditory priming showed no significant change from 2 years to college age. Auditory word priming in preschoolers was also robust against manipulations of the task during the initial presentation of the study words. As found previously for adults (Church & Schacter, 1994; Schacter & Church, 1992), the magnitude of auditory priming for 3-year-olds did not depend on whether the encoding task encouraged them to access the referent of each studied word. An object-choice task also revealed robust repetition priming in 18-month-olds (Fisher, Church, & Hunt, in

preparation): Infants were more likely to identify a familiar spoken word and correctly choose its referent if the word had been heard just twice in a prior listening phase. Other recent results yield evidence that auditory word priming in preschoolers permits abstraction across tokens (Fisher & Church, in press): 3-year-olds who heard words spoken in isolation or in short sentences later more accurately identified studied as opposed to new words presented in long sentences.

The similarity of auditory word priming in preschoolers and adults, not only in magnitude but also in its characteristic insensitivity to manipulations of the encoding task, strongly supports the hypothesis that the same learning mechanism is at work in both groups. The experiments reported in this paper test whether the same kind of perceptual facilitation can be demonstrated for preschoolers' identification of entirely novel word forms (Experiment 1), and ask whether auditory word priming in preschoolers shows the encoding of both abstract and specific components that we argued is essential for learning to identify the sound patterns of words (Experiments 2, 3, and 4).

### Experiment 1

In this study, 2.5-year-olds listened to one-syllable nonsense words (e.g., "yeeg" [jig], "lell" [lɛl]) in an initial study phase, participated in a distractor task, and then tried to identify and repeat a set of one-syllable nonsense words. Half of the test nonwords had been heard in the initial study phase, and half were new. Priming was measured by comparing performance on studied nonwords with performance on new items. The dependent variable was the proportion of phonemes in the target nonwords that were correctly repeated at test. No referent was provided for the nonwords at study or test; thus we asked whether minimal exposure to the sounds of new words increased young children's ability to perceptually identify their sound patterns.

As in previous studies (Church & Fisher, 1998; see also Gerken, Landau, & Remez, 1990), we used elicited imitation to explore what children can perceive about speech. Accurate repetition requires the perceptual identification of the stimulus items at test. Any advantage for

items heard in the study phase must be based on long-term storage of perceptual information about the sounds of the nonwords.

## Method

### Participants

Twenty-four 2.5-year-old children (28 to 31 months;  $M = 29.4$  months; 12 boys and 12 girls) participated in the experiment. The children were native speakers of English whose parents reported they had no exposure to a second language. Children received a small toy or book for their participation. Families were recruited for all experiments from a participant file drawn from birth announcements in the local newspaper. An additional 29 children (28-32 months;  $M = 29.3$  months) were tested but not included in the final data set because they did not produce enough relevant responses<sup>1</sup>. The children's parents took home and completed a short form of the MacArthur Communicative Development Inventory (Fenson et al., 2000) designed to assess the productive vocabulary of children 16-30 months old. Children who did not complete the task (27 forms returned, mean productive vocabulary score = 58 (of 100 possible),  $SD = 21.14$ ) had significantly lower vocabulary scores than those who did complete the experiment (23 forms returned,  $M = 79$ ,  $SD = 18.72$ ;  $t(48) = 3.69$ ,  $p < .001$ ). Given the productive demands of the repetition task, children with lower productive vocabularies had more difficulty completing the experiment.

### Stimuli

The stimuli for this experiment were 32 single syllable consonant-vowel-consonant (CVC) nonwords, adapted from a stimulus set used by Jusczyk et al. (1994). Jusczyk et al.'s stimuli included 240 nonwords consistent with the phonological rules of English (e.g., "biss" [bɪs], "cheedge" [tʃidʒ]).

Thirty-six pairs of nonwords matched on initial consonant were chosen from the Jusczyk et al. stimulus set, or in a few cases generated by the authors for this experiment. The nonwords were recorded by a female native speaker of English, using the SoundEdit16 software package on a Macintosh computer. To lessen the probability that children would mistake the nonwords

for particular English words, the stimuli were pre-tested with 24 adult subjects. The adults wrote down 3 English words that sounded similar to each of the 72 nonwords, and then rated the similarity of each English word they had generated to the nonword on a 7-point scale (1 = very similar sounding; 7 = not similar). Nonword pairs were discarded at this point if a particular English word was given a rating of "1" (very similar) by 4 or more adult subjects.

This step resulted in a set of 16 nonword pairs. The pairs were then divided into two lists, yielding two study lists with the same set of initial consonants. Stimulus lists are given in the Appendix for all three experiments.

The nonwords were recorded onto two study tapes of 16 nonwords each, and one test tape of all 32 items. To ensure that the children would hear all the nonwords despite occasional lapses of attention, each nonword was recorded twice onto the study tapes. Nonwords were separated by approximately 1 second of silence. The nonwords were randomly ordered, with the following constraints: On each study tape, all 16 items were played once before the list was repeated in a different random order. The test list order had no more than 3 items from the same study list in a row, and items from the two lists were equally represented in each quarter of the test list (e.g., 4 of the first 8 items were from study list A and 4 were from B).

A tape deck, amplifier, and speakers were used to present the stimuli to the children. A toy robot, "cookies" made out of foam, and a box with an opening in the top were used to make the experiment a game for the children. A small tape recorder with an external microphone was used to record the participants' responses.

### Procedures

The children were brought to the lab by their parents, and tested individually. Each session began with a 10 to 15 minute warm-up period to accustom the child to the lab and experimenter. After this play session, the experimenter brought the child into the testing room. The child sat in front of a table on which stood a speaker, a toy robot, a box containing toy cookies, and a box for putting the cookies into. The experimenter sat next to the child where he

or she could control presentation of the stimuli. If the parent accompanied the child into the testing room, he or she was asked to sit to the side and not to talk during the session.

In the study phase, the child was told that a robot would say "funny robot words." On each study trial, the experimenter prompted the child to listen, asked the child "Did that sound silly?" and prompted the child to reward the robot with a cookie. Study items were presented at a comfortable listening level (approximately 64-70 dB, as measured by a hand-held sound pressure meter at the position of the child's head).

The study phase was followed by a brief distractor task lasting at least 2 minutes. During this period the experimenter and child counted the cookies given to the robot, and played with the robot.

During the test phase the child heard all 32 nonwords once and tried to repeat them. The test words were presented at a much lower amplitude (approximately 54-60 dB), to increase the perceptual demands of the task. On each trial the experimenter prompted the child to listen, played the item, and prompted the child to repeat. The test phase began with four practice trials; two were nonwords spoken by the experimenter ("nowtch" [nautʃ] and "bawth" [bɔθ]) and two were recorded nonwords presented as in the test trials ("deece" [dis] and "yan" [jæn]). The test session was recorded for later transcription and coding.

### Coding

Children's attempts to imitate were broadly transcribed in the International Phonetic Alphabet (IPA) from audiotapes of the sessions, and each response was given an accuracy score equal to the proportion of target phonemes (out of 3) accurately repeated. The first-pass transcription and coding was checked by a second listener, and a third listener mediated all disagreements. The first listener played all three phases of the experiment and recorded experimenter errors, parent interference, or repetitions of the nonwords in the study phase; thus the first listener was not blind to the child's condition. The second and third transcribers listened only to the test phase and therefore were unaware of which study list the child had heard.

Deviations from the target phoneme were not scored as inaccurate if the child mispronounced the target phoneme in the same way in two or more of the target nonwords. These mispronunciations were considered consistent errors and were scored as accurate repetitions. This rule was designed to permit some ordinary childhood mispronunciation (e.g., all initial [k]'s produced as [t]'s), while retaining a fairly strict criterion of accuracy. Failures to respond and irrelevant responses (e.g., responses involving the robot or cookies, “I don’t know”) received accuracy scores of 0. Fifty-two trials (6.77% of the data) were missing because of experimenter error (27), parental prompting (2), because the child repeated the item during the study phase (22), or because the response was not picked up well by the microphone and could not be coded (1).

All data were independently transcribed by a fourth listener (also blind to condition), and these transcripts were compared to the result of the three-pass transcription procedure. The two transcriptions were in agreement for 80% of phonemes transcribed.

### Results

Children more accurately identified and repeated nonwords that they had heard in the study phase ( $M = .577$ ,  $SE = .045$ ) than new nonwords ( $M = .529$ ,  $SE = .048$ ). Planned, one-tailed comparisons conducted by subjects and by items revealed a significant effect of Study on repetition accuracy ( $t_1(23) = 2.49$ ,  $p < .01$ ;  $t_2(31) = 1.97$ ,  $p < .05$ ). The advantage for studied items remained significant when the reliability coder's independent transcriptions were used as the basis for scoring ( $t_1(23) = 2.77$ ,  $p < .01$ ;  $t_2(31) = 2.21$ ,  $p < .05$ ).

The children's responses were distributed across all accuracy levels: 29% of responses included all three target phonemes, 27% matched 2 target phonemes, 15% were accurate for only 1 phoneme, and 23% shared no phonemes with the target (including trials in which the child made no response). Children's vocabulary scores were positively correlated with their baseline repetition accuracy (Pearson  $R = .47$ ,  $n = 23^2$ ,  $p < .05$ ), but not at all with the magnitude of the priming effect: Priming scores (differences in accuracy between studied and new items) were not correlated with vocabulary scores (Pearson  $R = .05$ ,  $n=23$ ).

## Discussion

In Experiment 1, 2.5-year-olds were significantly better able to identify and repeat CVC nonwords that they had heard just twice in the study phase, several minutes before. As predicted, upon this minimal initial exposure, children created representations of the sound patterns of brand-new spoken words that facilitated their later identification of the same items.

Findings of robust long-term priming for nonwords provide more evidence for the perceptual nature of the learning underlying auditory word priming. In the current data, the facilitation for studied nonwords can only be attributed to the simple experience of hearing the words in the study phase. The nonwords were presented in a non-referential context, so no semantic representation unique to each item was likely to be created. In addition, children only very infrequently repeated words in the study phase, and the few trials on which they did were dropped from the analysis (see Methods); thus speaking practice is not implicated in the priming effect we measured. As previously found for adults (e.g., Goldinger, 1998), a lasting representation of new sound patterns is created when 2.5-year-olds simply listen to unfamiliar words. These new representations become part of the continually adapting system for identifying spoken words.

With these results in hand, we can go on to examine the nature of these rapidly-formed representations of new words. We argued in the Introduction that a crucial property of auditory word priming in adults was its combination of abstract and specific components. Can a similar combination of abstract and specific components be found in young children's representations of new spoken words? Experiments 2 and 3 provide a first test of this question. In both studies, children listened to CVC nonsense syllables presented in two-syllable sequences (CVC.CVC ; e.g., "sudyeeg" [sʌdjig], "deecegibe" [disgaɪb]).

Experiment 2 tested for abstract priming that spans a change in phonetic context. Children heard novel bisyllables in an initial study phase, and then repeated bisyllables in a test phase. Half of the test items were composed of two studied syllables, and half were composed of syllables not heard in the study phase (New syllables). All studied syllables were re-paired with

different studied syllables at test, to create Changed Context items. For example, a child who heard [disgaɪb] and [tʌljɛv] in the study phase would hear [tʌlgaɪb] and [disjɛv] in the test phase. The study and test items were recorded as naturally coarticulated bisyllables; thus both acoustic token and adjacent context changed from study to test. An advantage for studied syllables in this task will indicate that children's rapidly formed representations of novel words contain components abstract enough to support identification of items across phonetic contexts and acoustic changes.

Experiment 3 tested for specificity effects, or reductions in priming due to a change in the way each item is presented. All syllables of the bisyllabic nonwords were heard at both study and test—thus all were studied syllables. Half of the items changed both token and context from study to test, while half retained the same syllable token and context from study to test. For example, a child might hear [tʌljɛv] and [disgaɪb] in the study phase, then hear the syllables in the same context ([tʌljɛv] and [disgaɪb] again) or a changed context ([tʌlgaɪb] and [disjɛv]) at test. Performance on Changed Context items was compared to Same Context items to examine the specificity of children's representations of the novel items. Experiments 2 and 3 were conducted with children somewhat older than the children in Experiment 1 (36 to 42 months old), since we anticipated that the bisyllabic nonwords would be significantly harder to identify and repeat in the same procedure.

This basic method for examining the representation of "context" by looking at the effects of repairing co-presented information has been used by a number of researchers to examine adult implicit memory for new associations between words in both the visual (e.g. Gabrieli, Keane, Zarella, & Poldrack, 1997; Graf & Schacter, 1987, 1989; Goshen-Gottstein & Moscovitch, 1995a, 1995b; Light et al., 1995; Moscovitch, Winocur, & McLachlan, 1986; Musen, Shimamura, & Squire, 1990; Schacter & Graf, 1986, 1989) and auditory domains (e.g. Poldrack & Church, 1997). The current studies differ from the prior auditory experiments in a crucial way. The prior experiments presented familiar words in various contexts (e.g., paper-apple vs. book-apple); and tested for associations between familiar meaningful words and the contexts in

which they appeared. In the current experiments, there is no reason for the children to treat the nonsense syllables as separate meaningful units. Unfamiliar bisyllables with a CVC.CVC pattern, recorded with first-syllable stress, could sound like novel monomorphemic words (as in biscuit or council) or compound words (as in nightgown or madcap). The current studies explore what children encode about syllables in bisyllabic contexts when the syllables they hear are unfamiliar, and therefore provide no firm information about the locations of word boundaries.

## Experiment 2

Experiment 2 seeks evidence for abstraction in rapidly-formed perceptual representations of novel words, by asking whether previous experience with a novel syllable recorded in a different context is better than no previous experience with that syllable.

## Methods

### Participants

Twenty-four 3-year-olds (36 to 42 months;  $M = 37.3$  months; 12 boys and 12 girls) participated in the experiment. The children were native speakers of English whose parents reported that they had no significant exposure to a second language. Children received a small toy or book for their participation. An additional 13 children (36-40 months;  $M = 37$  months) were tested but were not included because they did not complete the task<sup>3</sup>.

### Stimuli

The bisyllabic nonwords for this study were created by combining the CVC nonwords selected for Experiment 1 with 32 new CVC nonwords chosen from the same source (Jusczyk et al., 1994). Five CVC nonwords from Experiment 1 were replaced with new syllables, to permit an examination of the effect of phonotactic frequency on repetition accuracy.

Each CVC syllable was recorded in two different pairings. These pairings were generated based on the following constraints: (1) Each bisyllable had one syllable of high phonotactic frequency and one of low phonotactic frequency. Half of the items had their high-frequency syllable in first position, and half had their high-frequency syllable in second position. For example, “biss” in [ʃaɪp.bɪs] is a syllable of high phonotactic frequency because many

English words share the onset [b] (as in ball), the vowel [ɪ] (as in big), or the coda [s] (as in boss). The biphone sequences in [bɪs] also occur in many English words (e.g., big, bin, kiss, miss). “Shipe” [ʃaɪp], on the other hand, has phonemes and biphone sequences which occur relatively rarely in those positions in English words. Following Jusczyk et al. (1994) and Vitevitch and Luce (1999), we calculated log-frequency-weighted probabilities that each phoneme and each biphone (CV or VC) occurred in its position in an on-line lexicon of nearly 20,000 English words (see Nusbaum, Pisoni, & Davis, 1984). The average summed phone probability was .1679 for high-frequency syllables and .0665 for low; the average summed biphone probability was .0091 for the high frequency syllables and .0009 for low. (2) Syllables were combined into 2-syllable sequences in which the first and second syllables shared no phonemes (e.g., [disgaɪb] was selected, but not [disjɪtʃ]), in which both syllables would have the same vowel). (3) The two syllables paired with each syllable also shared no phonemes (e.g., [disgaɪb] and [disjev] were selected, but not [disgaɪb] and [disgaʊr], because the two second syllables paired with [dis] would have shared an onset consonant). Thus the two syllables of each item, and the two recording contexts for each syllable, were as different from each other as possible.

The bisyllabic nonwords were recorded with primary stress on the first syllable, resulting in nonsense items with the phonological structure of English words like "bathtub" or "council." The less stressed second syllable always had a full vowel. This stress pattern was adopted to encourage perception of these items as bisyllabic words: Since even infants use the familiar initial-stress pattern of English words to group syllables into word-like units (e.g., Mattys et al., 1999; Morgan & Saffran, 1995), we judged that this sound pattern would encourage perception of these items as possible bisyllabic words.

The single-syllable nonwords from Experiment 1 (with 5 items replaced) were always the second, less stressed syllables of the new bisyllabic items; the 32 added nonsense syllables were always the first, more stressed syllables. Although both syllables of each bisyllabic nonword changed pairing from study to test, we expected that more robust priming would be found on the

second syllable. We anticipated that baseline repetition accuracy would be lower for the second syllable than for the first, since less stressed syllables are quieter and shorter than stressed syllables, and thus typically harder to identify (e.g., Echols & Newport, 1992; Echols, 1993). The magnitude of priming effects depends on baseline accuracy (e.g., Chapman, Chapman, Curran, & Miller, 1994). Our experience with these tasks suggests that larger priming effects are generally found when baseline accuracy is lower.

The resulting bisyllabic nonwords were divided into four study lists of 16 items. Each syllable appeared on two of the lists, in different pairings (e.g., [disgaɪb] and [tʌljev] on one list; [tʌlgaɪb] and [disjev] on another). Two test lists of 32 items were created. The study and test lists were combined so that each bisyllable was a Changed Context item for half of the children and a New item for the other half of the children. Study and test lists were balanced for the position (final/initial) of the high phonotactic frequency syllable in each bisyllable, and roughly balanced for the onset consonant of the second syllables.

The nonwords were recorded as in Experiment 1, by a female native speaker. Nonword tokens were re-recorded if the authors judged that they had been mispronounced in any respect (i.e., did not yield the intended phonemic transcription). The final lists were recorded onto audiotapes. As in Experiment 1, each stimulus token was repeated on the study tapes; in this case, each nonword was played twice in a row, separated by approximately 500 ms of silence. Different nonwords were separated by 1 s of silence. On the test tapes, each item appeared only once, with 1 s of silence between items, to allow the experimenter to pause the tape between test items.

### Procedure

The procedures were the same in general outline as those used in Experiment 1, with several modifications designed to make the task more interesting for the children, and to reduce the likelihood that they would repeat items in the study phase: (1) During the Study phase, the children simply listened to a study list while watching a brief portion of a Bugs Bunny™ cartoon video with the sound turned off. The nonwords were played from a speaker placed on top of the

video monitor, and the children sat in a chair pulled up to a table with a hole cut in its front edge, designed to keep the children the same distance from the speaker. The tape was not paused between items during the study phase, and the children made no judgments about the nonwords. The Study lists were presented at a comfortable listening level. (2) During the Distractor phase, the experimenter helped the child to work a puzzle for at least 2 minutes. (3) During the Test phase the children heard 32 bisyllabic nonwords: 16 were Changed Context items, and 16 were entirely New items. At test the volume was reduced so that the nonwords were very quiet (approximately 52-56 dB). The children were told to listen and repeat each nonsense word. They were warned that the words would be hard to hear and that they would have to listen carefully. Children were given stickers as rewards for responding. The test phase began with two practice trials; on these trials the experimenter prompted the child to respond, and demonstrated the task if the child did not respond at first. All responses were recorded for later transcription and coding.

### Coding

The data were transcribed and coded using the same three-pass system described for Experiment 1. A randomly chosen 25% of the data were independently transcribed and coded by a fourth listener. The fourth listener's transcriptions were in agreement with the output of the three-pass transcription system for 74% of phonemes transcribed.

The transcription resulting from the three-pass system was coded in the following manner. Each phoneme received a score of 1 or 0: 1 if the child repeated the phoneme correctly or repeated it incorrectly but did so consistently (e.g., incorrectly substituting [w] for [l] in two or more test items), and 0 if the child repeated the phoneme incorrectly. If the participant made no response, or an irrelevant response, all six phonemes of that test trial received a score of 0. The vast majority of children's repetition attempts had the [CVC.CVC] structure of the stimulus items; responses were coded in order such that the first phoneme of the response was assumed to be an attempt at the first phoneme of the target. Children occasionally produced only one consonant at the syllable boundary; in these cases the medial consonant was considered an

attempt at the initial phoneme of the second syllable (thus, [CV.CVC]). This coding rule was designed to give a consistent way to compare repetition attempts to targets, without attempting to base our decisions on judgments of phonetic similarity. This rule was chosen because it best fit listeners' intuitions about the appropriate syllabification of the responses, and to honor children's tendency to omit syllable-final consonants more often than onset consonants in spontaneous speech (e.g., Vihman, 1996). The accuracy score for each syllable was the proportion of its three phonemes that were repeated accurately.

The children repeated none of the items in the study phase; 29 trials (3.78% of the data) were missing because the response was not recorded clearly enough to be coded (19), the child or parent talked while the test word was played (5), or because of experimenter error (5).

### Results

As shown in Table 1, children more accurately identified nonsense syllables that changed context from study to test than syllables that were entirely new at test. As predicted, baseline accuracy was lower for the second syllable, and the priming effect was correspondingly larger on second syllables. Two (First vs. Second syllable) by 2 (Changed Context Studied vs. New) ANOVAs conducted by subjects and by items revealed an effect of syllable ( $F_1(1,23) = 26.48, p < .001$ ;  $F_2(1, 63) = 12.27, p < .001$ ), reflecting the lower accuracy for second syllables, and an effect of study, significant by subjects but not by items ( $F_1(1,23) = 4.88, p < .05$ ;  $F_2(1,63) = 2.64, p = .109$ ), showing an advantage for changed-context studied over new syllables. The interaction between these two factors approached significance ( $F_1(1,23) = 2.80, p = .108$ ;  $F_2(1,63) = 2.50, p = .119$ ), reflecting the anticipated trend for the priming effect to be more robust on the second syllables, with their lower baseline accuracy. Planned comparisons revealed significant priming for the second syllables ( $t_1(23) = 2.96, p < .01$ ;  $t_2(63) = 2.22, p < .05$ ) but not for the first syllables (both  $t$ 's  $< 1$ ). The same pattern was found when the entire bisyllable was taken as the unit of analysis rather than the component syllables (Changed Context  $M = .545, SE = .031$ ; New  $M = .505, SE = .027$ ;  $t_1(23) = 2.21, p < .05$ ;  $t_2(63) = 1.62, p = .055$ ).

----- Insert Table 1 about here -----

As in Experiment 1, children's responses were broadly distributed across accuracy levels: The modal response matched the bisyllabic target on 5 of 6 phonemes (21% of responses), but substantial proportions of responses matched the target on 6 (9%), 4 (17%), 3 (16%), 2 (16%), 1 (9%) or no phonemes (12%).

High phonotactic frequency syllables ( $M = .535$ ,  $SE = .035$ ) also tended to be more accurately repeated than low phonotactic frequency syllables ( $M = .516$ ,  $SE = .022$ ). This trend was small, however, and a 2 (First vs. Second syllable) by 2 (High vs. Low frequency syllable) ANOVA conducted by subjects<sup>4</sup> revealed only a main effect of syllable ( $F_1(1, 23) = 27.27$ ,  $p < .001$ ), and no effect of phonotactic frequency ( $F < 1$ ), or interaction of syllable and frequency ( $F_1(1,23) = 1.79$ , *ns*).

### Discussion

The results of Experiment 2 tell us that children's rapidly formed representations of new spoken words can be used abstractly to facilitate identification of similar but not identical items, spanning a change in phonetic context and the associated change in acoustic token. Children more accurately repeated syllables they had heard just twice in the study phase; this was true even though all studied syllables changed syllable token and context from study to test. For preschoolers as for adults, the perceptual learning underlying auditory word priming permits abstraction across acoustic and context variability, even for brand-new items.

Experiment 2 yielded no strong evidence for an effect of phonotactic frequency on the identification and repetition of new word forms. We saw only a small trend toward greater accuracy for syllables of high phonotactic frequency.

### Experiment 3

The results of Experiment 2 demonstrated that auditory word priming in preschoolers reveals the kind of rapid abstraction over variation among tokens of the same sound pattern that would be required to learn to identify words in ordinary connected speech. However, learners must not only abstract across variation to identify words, but also retain specific information about acoustic variability and the contexts that explain some of that variability.

Do children, like adults, rapidly encode representations of spoken words that retain specific details of the original presentation? In Experiment 3, all test items were composed of syllables the children had heard in the study phase. Half of the test items were the same bisyllabic tokens heard in the study phase (Same Context items), while half were Changed Context items. An advantage for Same Context items would show that rapidly-formed representations of novel syllables retain specific information relevant to the context in which each syllable first appeared.

### Method

#### Participants

Twenty-four 3-year-old children (36.4 to 41.2 months;  $M = 37.6$  months; 12 boys and 12 girls) participated in the experiment. The children were monolingual native speakers of English. Children received a small toy or book for their participation. An additional 11 children (36 to 39 months;  $M = 37.5$  months) were tested but not included in the final data set because they did not produce enough relevant responses.

#### Stimuli

The 64 bisyllabic nonwords used for Experiment 2 were organized into new study and test lists for Experiment 3. The nonwords were divided into two study lists of 32 items. Each syllable appeared on both lists, but in different pairings (e.g., [disgarb] and [tʌljev] on one list; [tʌlgarb] and [disjev] on the other). The same nonwords were divided into two test lists. Study and test lists were combined so that each bisyllable was a Same Context item for half of the children and a Changed Context item for the other half of the children. As in Experiment 2, study and test lists were balanced for the position (initial/final) of the high phonotactic frequency syllable in each pair, and the second syllables were roughly balanced across lists for onset consonant.

We used the same 64 recorded tokens used in Experiment 2. The study and test lists were recorded onto audiotapes as in Experiment 2. On the study tapes, each item appeared twice in a row, with approximately 500 ms between repetitions, and each nonword separated from the next

nonword by 1s of silence. On the test tapes, each bisyllabic nonword appeared once, with 1s of silence between nonwords.

### Procedure

The procedure was identical to the procedure for Experiment 2, with the exception that the children heard 32 bisyllabic nonwords during the study phase. Children listened to the nonwords as they watched a silent cartoon in the study phase. Next they worked on a puzzle during the distractor phase. Finally, they attempted to repeat 32 quietly presented nonwords in the test phase.

### Coding

The children's repetition attempts were transcribed and coded as in Experiment 2. A randomly chosen 25% of the data were independently transcribed and coded by a fourth listener, and reliability was calculated between the three-pass transcriptions and the fourth assistant's transcriptions. The two were in agreement for 71% of transcribed phonemes.

The children repeated none of the items in the study phase; 15 trials (1.95% of the data) were missing because the response was not recorded clearly enough to be coded (9), the child talked while the item was played (2), because of experimenter error (3) or parental coaching (1).

### Results

As shown in Table 2, children more accurately repeated both first and second syllables at test if they were heard in the Same Context as at study, and were less accurate on Changed Context items. As expected, baseline accuracy was lower for the relatively unstressed second syllables, and the Same Context advantage was correspondingly larger for second syllables. These patterns were tested in 2 (First vs. Second syllable) by 2 (Same vs. Changed Context) ANOVAs conducted both by subjects and by items. These analyses revealed an effect of context change ( $F_1(1,23) = 5.51, p < .05$ ;  $F_2(1,63) = 3.59, p = .063$ ), with no significant effect of syllable ( $F_1(1,23) = 1.65, ns$ ;  $F_2(1,63) = 1.32, ns$ ) or interaction of context change and syllable ( $F_1(1,23) = 1.32, ns$ ;  $F_2(1,63) < 1$ ). Planned, one-tailed comparisons revealed a significant Same Context advantage for second syllables ( $t_1(23) = 3.05, p < .01$ ;  $t_2(63) = 1.97, p < .05$ ) but not for first

syllables ( $t$ 's < 1). The same pattern was found when the bisyllable was taken as the unit of analysis rather than the component syllables (Same Context  $M = .682$ ,  $SE = .023$ ; Changed Context  $M = .647$ ,  $SE = .025$ ;  $t_1(23) = 2.36$ ,  $p < .05$ ;  $t_2(63) = 1.81$ ,  $p < .05$ ).

----- Insert Table 2 about here -----

Children's responses were again broadly distributed across accuracy levels: The modal response matched the bisyllabic target on 5 of 6 phonemes (27% of responses), but substantial proportions of responses matched the target on 6 (19%), 4 (21%), 3 (16%), 2 (8%), 1 (4%) or no phonemes (6%).

As shown in Table 3, high phonotactic frequency syllables tended to be more accurately repeated than low phonotactic frequency syllables. A 2 (First vs. Second syllable) by 2 (High vs. Low phonotactic frequency) ANOVA conducted by subjects<sup>5</sup> revealed a main effect of phonotactic frequency ( $F(1,23) = 29.63$ ,  $p < .001$ ), and no main effect of syllable ( $F(1,23) = 1.54$ ,  $ns$ ). The effect of frequency did not interact with the effect of syllable ( $F < 1$ ). The effect of phonotactic frequency was significant by paired, 2-tailed  $t$ -tests for both the first and second syllables (1st syllable:  $t(23) = 4.28$ ,  $p < .001$ ; 2nd syllable:  $t(23) = 2.48$ ,  $p < .05$ ).

----- Insert Table 3 about here -----

### Discussion

The results of Experiment 3 demonstrate considerable specificity in 3-year-olds' rapidly formed representations of new spoken words. Children heard novel bisyllabic items just twice in the study phase. Based on this minimal exposure, they were better able to identify and repeat quietly presented syllables in the test phase if the same tokens of those syllables were heard in the same recorded context than if new tokens were presented in a new context. The advantage for same-context items was reflected in the means for both the first and second syllables, but the effect was most clearly seen on the second, less stressed syllable. This is the first evidence in preschoolers of the kind of specific encoding of spoken words that has been found in many recent studies with adults (e.g., Church & Schacter, 1994; Goldinger, 1996, 1998; Schacter & Church, 1992).

The data also revealed an advantage for syllables of high phonotactic frequency. In agreement with many studies of both infant and adult perception of new words (e.g., Mattys et al., 1999; Vitevitch & Luce, 1999; Jusczyk et al., 1994), 3-year-olds more accurately repeated syllables whose component sound sequences were of high frequency in the English lexicon. Gathercole (1995) has found that preschoolers more accurately repeat long nonwords that are rated as highly wordlike by adults, and that have syllables reminiscent of English morphology (e.g., sladding, doppelate). Such findings have been interpreted as evidence of lexical effects in the perception or production of new words. A similar account emerges from the adult literature: Vitevitch and Luce (1999) argue that when trying to identify nonwords, representations of familiar sequences smaller than a word are activated, and support identification of similar sequences in new arrangements.

The phonotactic frequency effect indicates abstraction across contexts in the use of long-term representations of the sound patterns of English words: By definition, since our items were nonwords, the familiar sound sequences of the high phonotactic syllables were heard in unprecedented larger contexts. Despite this context change, frequent experience with sublexical sequences made it easier to hear and repeat new words containing them. In the current task, greater accuracy on high phonotactic frequency syllables could reflect the effect of practice hearing English words, saying them, or both. The effect of a token and context change from study to test, however, could only be based on listening experience: Children did not repeat the items in the study phase, so could only profit from the experience of hearing them.

#### Experiment 4

What specific features of the bisyllabic nonwords did the children encode? Based upon Experiment 3, there are two classes of possibilities. Changed Context syllables differed in both adjacent context and recorded token from study to test. The context change itself, or changes in syllable token, or both, could have caused the reduction in priming for changed-context items. We argued in the Introduction that both kinds of information should be needed in the representations that support spoken word identification. To learn how speech sounds are

affected by various contexts in a specific language, children must create detailed and context-sensitive representations of spoken words.

To begin to explore the two components of the specificity effect measured in Experiment 3, we need to separate the influences of token change and context change. Experiment 4 was designed to isolate the effect of a token change; other studies, described in the General Discussion, are designed to isolate the effect of a context change.

Token changes due to coarticulation (e.g., [bɪs] taken from [ʃaɪpɪs] vs. [dʒeθɪs]) are small sound changes. However, previous data from adult listeners suggest that such token changes can cause significant reductions in priming. Poldrack and Church (1997) had adults listen to semantically unrelated words recorded as coarticulated pairs (e.g., "client-weather", "paper-apple"). At test, adults more accurately identified low-pass filtered versions of these words if they appeared in the same context (e.g., "paper-apple") than in a changed context (e.g., "client-apple"). Changed-context items were more accurately identified than entirely new items. Thus the adults showed the same kind of context-change reduction in priming that the preschoolers showed in Experiment 3. In subsequent experiments, Poldrack and Church found that the token change alone produced a reduction in priming of a magnitude similar to the effect found in their first study. When words were excised from their recording contexts, subjects more accurately identified a filtered version of the same word token they had heard in the study phase than the token recorded in the other context. The analogy between our studies and Poldrack and Church's suggests that a token change is likely to be at least partly responsible for the Same-Context advantage found in Experiment 3.

In Experiment 4, three-year-olds heard the same study lists as in Experiment 3. At test, however, the children heard only the second syllable of each test item, spliced out of its bisyllabic context. As in Experiment 3, all test items were syllables the children had heard in the study phase. Half of the test syllables were the same syllable tokens heard in the study phase, spliced out of Same Context bisyllables (Same Token items), while half were different syllable tokens, spliced out of the Changed Context bisyllables (Changed Token items). An advantage

for Same Token items would show that children's representations of novel syllables preserved information particular to the token which they originally heard, independent of whether the syllable was presented in the same context.

We expected the advantage for same-token items in Experiment 4 to be somewhat harder to detect than the advantage for same-token/same-context items in Experiment 3. Any contribution that the context change itself made to the specificity effect of Experiment 3 is missing in Experiment 4; thus the token-specificity effect of Experiment 4 is likely to be a smaller effect. For this reason, Experiment 4 had 32 subjects rather than 24.

### Method

#### Participants

Thirty-two 3-year-old children (36.2 to 41.2 months;  $M = 38.7$  months; 16 girls and 16 boys) participated in the experiment. The children were monolingual native speakers of English. Children received a small toy or book for their participation. An additional 6 children (36.7 to 38.6 months;  $M = 37.7$  months) were tested but not included in the final data set because they did not produce enough relevant responses.

#### Stimuli and Procedures

Experiment 4 was identical to Experiment 3 except that the second syllables of the original bisyllabic items were presented in isolation at test. Test items from Experiment 3 were altered by splicing the second syllable out of its context bisyllable using a wave-form editor. The isolated second syllables of these items were divided into two test lists; the study and test lists were combined so that each test token was a Same Token item for half of the children and a Changed Token item for the other half of the children. The study and test lists were recorded onto a compact disc for playback. In the study lists, each item appeared twice in sequence, with approximately 500 ms between repetitions, and each nonword separated from the next nonword by 1s of silence. In the test lists, each nonsense syllable appeared only once, with 1s of silence between items, to permit the experimenter to pause playback and prompt the child to repeat.

#### Coding

The children's repetition attempts were transcribed and coded as in Experiments 2 and 3. A randomly chosen 25% of the data were independently transcribed and coded by a fourth listener, and reliability was calculated between the three-pass transcriptions and the fourth assistant's transcriptions. The two were in agreement for 89% of transcribed phonemes.

Thirteen trials (1.3% of the data) were missing because of experimenter error (11) or parental coaching (2).

### Results

Children more accurately identified the nonsense syllables at test if they heard the Same Token previously heard at study ( $M = .676$ ,  $SE = .021$ ) than if they heard a Changed Token ( $M = .632$ ,  $SE = .021$ ). Planned, one-tailed t-tests conducted both by subjects and by items revealed a significant advantage for Same Token items ( $t_1(31) = 1.81$ ,  $p < .05$ ;  $t_2(31) = 2.50$ ,  $p < .01$ ). The advantage for Same over Changed Token items in Experiment 4 (.043) was only slightly smaller than the advantage for the Same over Changed Context items in Experiment 3 (.050), in which both context and token were preserved or changed from study to test.

The majority of the children's responses shared 2 (39%) or 3 (34%) phonemes with the target; 17% matched the target on one phoneme, and 10% shared no phonemes with the target.

As in Experiment 3, high phonotactic frequency syllables tended to be more accurately repeated than low phonotactic frequency syllables (High  $M = .681$ ,  $SE = .018$ ; Low  $M = .627$ ,  $SE = .019$ ). The advantage for syllables of high phonotactic frequency was significant in planned, 1-tailed t-tests by subjects but not by items ( $t_1(31) = 3.36$ ,  $p < .001$ ;  $t_2(30) = 1.51$ ,  $p = .071$ ).

### Discussion

In Experiment 3 we found that children's rapidly-formed representations of new spoken words included quite detailed information about the original item — either acoustic-phonetic information about syllable tokens, information about a syllable's adjacent context, or both. The results of Experiment 4 make clear that a significant part of the specificity effect found in Experiment 3 was due to the encoding of token-specific information about the sounds of

syllables. Three-year-olds were more accurate when asked to identify and repeat nonsense syllables which were the same tokens heard in the study phase, than different tokens of the same syllables. This finding is consistent with many results from the adult literature on auditory priming: The advantage of prior listening experience is reduced by study-to-test changes in the realization of the same familiar or unfamiliar word. As predicted, young children routinely encode and retain token-specific information about the sound patterns of novel words.

We again found an advantage for syllables of high phonotactic frequency. As in Experiment 3, 3-year-olds more accurately repeated syllables whose component sound sequences were of high frequency in the English lexicon.

### General Discussion

Acquiring a language involves learning an enormous amount about the sound patterns of the native language. This learning takes place at many levels of generality: We learn to identify the sound patterns of individual words, but also become sensitive to regularities in how sounds are sequenced within words. Listeners learn the typical prosodic patterns of words, phrases, and sentences in a language, and develop language-specific quantitative estimates of the many acoustic-phonetic parameters that differ across languages, from the timing of components of voiced versus voiceless consonants to the degree of utterance-final vowel lengthening, and the many ways in which the details of speech sounds depend on their contexts.

Learning these sound patterns requires a learning mechanism that can encode both phonetic context information and acoustic/phonetic detail. Prior research has uncovered a powerful learning mechanism in adults which seems to fit these requirements. Research in auditory word priming shows that adults continually add new perceptual information to the memory systems subserving word identification (e.g., Church, 1995; Church & Schacter, 1994; Goldinger, 1996, 1998; Pilotti et al., 2000; Schacter & Church, 1992; Sheffert, 1998; Sommers, 1999). Explorations of auditory implicit memory reveal that the underlying mechanisms have many of the same properties that we argued are needed to acquire appropriately detailed and flexible representations of the sounds of words. Our prior work provides initial support for the

claim that the same learning mechanism operates continuously throughout development (e.g., Church & Fisher, 1998).

The current studies add two important elements to this argument. First, we found auditory priming for nonwords in young preschoolers. In all four experiments, children more accurately repeated nonwords that were identical (Experiments 1, 3, and 4) or similar (Experiment 2) to items played in the study phase, than items that were new (Experiments 1 and 2) or less similar (Experiments 3 and 4) to the study items. Simply listening to novel words, each presented twice without unique referential context, made it easier for 2.5- and 3-year-olds to hear and repeat the same or similar nonwords later. These results provide new support for the perceptual nature of the facilitation we measured. As for adults (e.g., Goldinger, 1998), hearing the sound patterns of novel words eases the later identification of those words.

Second, Experiments 2, 3, and 4 provide the first evidence that auditory word priming in preschoolers creates representations of spoken words that combine abstract and token-specific components. Upon brief exposure to a set of entirely novel bisyllabic words, 3-year-olds rapidly encoded representations of the syllables which could be used abstractly to identify new tokens of the same syllables in different contexts. Thus, in Experiment 2, a Changed Context repetition was better than none. Based on the same minimal exposure, however, other 3-year-olds showed that they had encoded representations of the novel syllables that included more specific components. In Experiment 3, children gained more facilitation for later identification and repetition if the nonwords remained the same (e.g., [ʃaɪpbɪs] → [ʃaɪpbɪs]) than if they changed both token and context from study to test (e.g., [ʃaɪpbɪs] → [dʒeθbɪs]).

Findings in the adult literature (Poldrack & Church, 1997) suggested that a token change was likely to be partly responsible for the context-change reduction in priming that we found in Experiment 3. The results of Experiment 4 confirmed this prediction: Children showed almost as great an advantage if the nonwords were the same tokens heard at study rather than different tokens recorded in different contexts, even though both Same and Changed tokens were excised from their original contexts (e.g., [bɪs] from [ʃaɪpbɪs] vs. [dʒeθbɪs]). Children rapidly encoded

representations of novel words that included token-specific details, yet were also able to use their representations abstractly to support the identification of similar syllables in new contexts (Experiment 2). A long series of studies documents a similar combination of abstraction and specificity in auditory priming in adults (e.g., Church & Schacter, 1994; Church et al., 1996; Goldinger, 1996, 1998; Pilotti et al., 2000; Schacter & Church, 1992; Sheffert, 1998; Sommers, 1999).

Did context play any role in the Same-Context advantage measured in Experiment 3? With clear presentation of the context word, Poldrack and Church (1997) found no reduction in priming when only the context changed. If the same word tokens were spliced into different contexts at study and test, the adults were no more accurate for same-context than for changed-context items<sup>6</sup>. Adult listeners in their tasks routinely encoded token-specific information about the sounds of words that affected later word identification, but did not appear to use perceptual information about the contexts in which words appeared to identify spoken words at test.

This finding need not suggest, however, that perceptual representations of the context itself can have played no role in our studies. The materials in these sets of studies were very different from those used by Poldrack and Church (1997). The children in our experiments listened to novel syllables presented in bisyllabic sequences, recorded with a stress pattern reminiscent of many English bisyllabic nouns. The adults in the Poldrack and Church experiments listened to familiar words in semantically odd sequences (e.g., "book-flamingo"). These words constituted familiar ready-made units within the sequences the adults heard. It may be that cues relevant to word segmentation, including the identification of familiar words, stress pattern, and phonotactic probabilities, influence whether, and how strongly, contextual information is used to influence word identification. Information about context might under some circumstances be encoded as relational or associated information, which should have a relatively weak effect on the hearer's ability to identify the same item again (e.g., Musen & Squire, 1993), but permit considerable abstraction across contexts. Under other circumstances an item-in-context might be represented as a more tightly integrated perceptual unit, permitting

less abstraction across contexts (e.g., Cohen & Eichenbaum, 1993). Components of auditory priming which are specific to a particular context may be easier to detect in cases where the boundaries of familiar words do not provide strong segmentation cues.

In order to isolate the influence of context from token-internal sound changes in repetition priming, studies currently underway test 3-year-olds with spliced-together versions of the bisyllabic stimuli used in Experiment 3. Preliminary findings using these materials have revealed small but consistent effects of a change in adjacent context, even though the recorded syllable token remains the same from study to test (Chambers, Fisher, & Church, 1999). These data suggest that children rapidly encode information about the context in which an unfamiliar syllable occurs—at least in bisyllabic items with greater stress on the first syllable. Further studies will need to manipulate cues relevant to word segmentation, to explore whether these cues mediate the appearance of abstract and context-specific priming effects, and thus whether they play a role in the creation of perceptual units in rapidly-formed representations of input speech.

The current findings also provide some further evidence that listeners of all ages are more accurate in identifying nonword sequences of high phonotactic probability. In Experiments 2, 3, and 4, children showed trends toward higher accuracy in identifying and repeating syllables of high rather than low phonotactic frequency. This effect was significant for Experiments 3 and 4. As found for infants in listening preference tasks (e.g., Jusczyk et al., 1994; Mattys et al., 1999) and for adults in shadowing tasks (e.g., Vitevitch & Luce, 1999), preschoolers are influenced by the frequency of subsyllabic patterns in speech. Syllables composed of sequences which are frequent in the lexicon (e.g., [bɪ] as in "bit" and "bin") are easier to identify and repeat than sequences which are less frequent in the lexicon (e.g., [gəʊ] as in "gown"). Such phenomena implicate a strong link between repeated perception or production of particular sound patterns in context, and subsequent processing of identical or similar items. This suggests an influential role in speech processing for the kind of learning mechanism explored here, combining detailed

records of experience identifying spoken words, and an ability to readily abstract across these details.

Our findings suggest considerable continuity across development in learning procedures which build long-term memory representations of the sounds of words. Preschoolers, like adults, possess a learning mechanism that creates and updates long-term representations of the sounds of words to reflect ongoing experience with language. This rapid perceptual learning facilitates later identification of the same or similar sound patterns. The representations laid down when children listen to speech include perceptual details specific to the syllable tokens originally presented, yet also can be used abstractly to support identification of similar but not identical patterns.

We would argue that many aspects of languages' sound systems could not exist without a learning mechanism with these properties. As reviewed in the Introduction, languages vary not only in the sound patterns that make up particular words (e.g., "dog" [dɔg] is a word in English but not in French), but also in their phonotactic rules (e.g., "ng" [ŋ] is a legal word onset in Vietnamese but not in English), in the phonetic realization of phonologically similar contrasts (e.g., Spanish & English differ in the exact voice-onset-time distributions signaling voiced and voiceless consonants), and in the vulnerability of particular segments to coarticulatory influences (e.g., vowels are lengthened before voiced stop consonants in English more than they are in Polish). In order to learn these sound-pattern regularities, learners must encode fairly detailed acoustic-phonetic information about speech sounds, and retain some information about the context in which they occurred.

We have presented data suggesting that preschoolers' rapidly-formed representations of spoken words combine the abstraction and specificity needed to learn such regularities. These data support the continuity of learning hypothesis—that the same perceptual learning mechanism which permits adults to adapt continually to new words, changing speakers and acoustic circumstances (e.g., Nusbaum & Goodman, 1994; Pisoni, 1992) also plays a role in the initial acquisition of the auditory lexicon. Further study of the nature of this learning mechanism, and

the kinds of perceptual details it readily encodes, will yield valuable new information about how the perceptual system for identifying spoken words adapts to experience during the course of language acquisition.

## References

- Althoff, R. R., & Cohen, N. J. (1999). Eye-movement-based memory effect: A reprocessing effect in face perception. Journal of Experimental Psychology: Learning, Memory, & Cognition, 25, Jul 1999, 997-1010.
- Beckman, M., & Pierrehumbert, J. (1986). Intonational structure in Japanese and English. Phonology Yearbook, 3, 255-309.
- Campbell, N. (1992). Segmental elasticity and timing in Japanese speech. In Y. Tokhura, E. Vatikiotis-Bateson, & Y. Sagisaka (Eds.), Speech perception, production, and linguistic structure (pp. 403-418). IOS Press.
- Carey, S., & Bartlett, E. (1978). Acquiring a single new word. Papers and Reports on Child Language Development, 15, 17-29.
- Chambers, K., Fisher, C., & Church, B. (1999). Context-sensitive long-term auditory word priming in 3-year-old children. Paper presented at the Annual Meeting of the Cognitive Development Society, Chapel Hill NC, October.
- Chapman, L. J., Chapman, J. P., Curran, T., E., & Miller, M. B. (1994). Do children and the elderly show heightened semantic priming? How to answer the question. Developmental Review, 14, 159-185.
- Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. Phonetica, 22, 129-159.
- Church, B. A. (1995). Perceptual specificity of auditory priming: Implicit memory for acoustic information. Unpublished doctoral dissertation, Harvard University, Cambridge, MA.
- Church, B. A., Dell, G., & Kania, E. (1996). Representing phonological information in memory: Evidence from auditory priming. Paper presented at the meeting of the Psychonomics

Society.

- Church, B. A., & Fisher, C. (1998). Long-term auditory word priming in preschoolers: Implicit memory support for language acquisition. Journal of Memory & Language, 39, 523-542.
- 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, & Cognition, 20, 521-533.
- Cohen, N. J., & Eichenbaum, H. E. (1993). Memory, amnesia, and the hippocampal system. Cambridge, MA: MIT Press.
- Crystal, T. H., & House, A. S. (1988). Segmental durations in connected-speech signals: Current results. Journal of the Acoustical Society of America, 83, 1553-1573.
- Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. Journal of Experimental Psychology: Human Perception and Performance, 14(1), 113-121.
- Dahan, D., & Brent, M. R. (1999). On the discovery of novel wordlike units from utterances: An artificial-language study with implications for native-language acquisition. Journal of Experimental Psychology: General, 128, 165-185.
- Dell, G.S., Reed, K.D., Adams, D.R., & Meyer, A.S. (2000). Speech errors, phonotactic constraints, and implicit learning: A study of the role of experience in language production. Journal of Experimental Psychology: Learning, Memory & Cognition, 26, 1355-1367.
- Echols, C. (1993). A perceptually-based model of children's earliest productions. Cognition, 46, 245-296.
- Echols, C. H., & Newport, E. L. (1992). The role of stress and position in determining first

- words. Language Acquisition, 2, 189-220.
- Farnetani, E. (1997). Coarticulation and connected speech processes. In W. J. Hardcastle & J. Laver (Eds.), The Handbook of Phonetic Sciences, (pp. 371-404). Oxford: Blackwell Publishers.
- Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur Communicative Development Inventories. Applied Psycholinguistics, 21, 95-115.
- Fisher, C., & Church, B. A. (in press). Implicit memory support for language acquisition. To appear in J. Weissenborn & B. Hoehle (Eds.), Approaches to bootstrapping: Phonological, syntactic, and neurophysiological aspects of early language acquisition. Amsterdam: Benjamins.
- Fisher, C., Church, B. A., & Hunt, C. (in preparation). Auditory word priming in 18-month-old infants.
- Fisher, C., Hall, G.D., Rakowitz, S., & Gleitman, L. R. (1994). When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. Lingua, 92, 333-375.
- Fisher, C., & Tokura, H. (1996a). Acoustic cues to grammatical structure: Cross-linguistic differences. Child Development, 67, 3192-3218.
- Fisher, C., & Tokura, H. (1996b). Prosody in speech to infants: Direct and indirect acoustic cues to syntactic structure. In J.L. Morgan & K. Demuth (Eds.), Signal to syntax: Bootstrapping from speech to syntax in early acquisition. Hillsdale, NJ: Erlbaum.
- Gabrieli, J. D. E., Keane, M. M., Zarella, M. M., & Poldrack, R. A. (1997). Preservation of implicit memory for new associations in global amnesia. Psychological Science, 8, 326-329.

- Gathercole, S. E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. Memory & Cognition, 23, 83-94.
- Gerken, L., Landau, B., & Remez, R. E. (1990). Function morphemes in young children's speech perception and production. Developmental Psychology, 26(2), 204-216.
- 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.
- Goodsitt, J. V., Morgan, J. L., & Kuhl, P. K. (1993). Perceptual strategies in prelingual speech segmentation. Journal of Child Language, 20, 229-252.
- Gordon, P. C. (1989). Context effects in recognizing syllable-final /z/ and /s/ in different phrasal positions. Journal of the Acoustical Society of America, 86, 1698-1707.
- Goshen-Gottstein, Y., & Moscovitch, M. (1995a). Repetition priming for newly formed and preexisting associations: Perceptual and conceptual influences. Journal of Experimental Psychology: Learning, Memory, & Cognition, 21, 1229-1248.
- Goshen-Gottstein, Y., & Moscovitch, M. (1995b). Repetition priming effects for newly formed associations are perceptually based: Evidence from shallow encoding and format specificity. Journal of Experimental Psychology: Learning, Memory, & Cognition, 21, 1249-1262.
- Gow, D. W., Jr., & Gordon, P. C. (1995). Lexical and prelexical influences on word segmentation: Evidence from priming. Journal of Experimental Psychology: Human Perception and Performance, 21, 344-359.

- Graf, P. & Schacter, D. L. (1987). Selective effects of interference on implicit and explicit memory for new associations. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13, 45-53.
- Graf, P., & Schacter, D. L. (1989). Unitization and grouping mediate dissociations in memory for new associations. Journal of Experimental Psychology: Learning, Memory, & Cognition, 15, 930-940.
- Harris, J. (1994). English sound structure. Oxford: Blackwell.
- Heibeck, T. H., & Markman, E. M. (1987). Word learning in children: An examination of fast mapping. Child Development, 58, 1021-1034.
- Jackson, A., & Morton, J. (1984). Facilitation of auditory word recognition. Memory & Cognition, 12, 568-574.
- Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. Journal of Experimental Psychology: General, 110, 306-340.
- Jusczyk, P. W. (1997). The discovery of spoken language. Cambridge, MA: MIT Press.
- Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. Journal of Memory & Language, 33, 630-645.
- Keating, P. (1985). Universal phonetics and the organization of grammars. In V. Fromkin (Ed.), Phonetic linguistics: Essays in honor of Peter Ladefoged (pp. 115-132). New York: Academic Press.
- Keating, P. A. (1990). Phonetic representations in a generative grammar. Journal of Phonetics, 18, 321-334.
- Klatt, D. H. (1976). Linguistic uses of segmental durations in English: Acoustic and perceptual evidence. Journal of the Acoustical Society of America, 59, 1208-1221.

- Klatt, D. H. (1980). Speech perception: A model of acoustic-phonetic analysis and lexical access. In R. A. Cole (Ed.), Perception and production of fluent speech (pp. 243-288). Hillsdale, NJ: Erlbaum.
- Kluender, K. R., Diehl, R. L., & Wright, B. A. (1988). Vowel-length differences before voiced and voiceless consonants: An auditory explanation. Journal of Phonetics, 16, 153-169.
- Light, L. L., La Voie, D., & Kennison, R. (1995). Repetition priming of nonwords in young and older adults. Journal of Experimental Psychology: Learning, Memory, & Cognition, 21, 327-346.
- Lively, S. E., Pisoni, D. B., & Goldinger, S. D. (1994). Spoken word recognition. In M. A. Gernsbacher (Ed.), Handbook of psycholinguistics (pp. 265-301). New York: Academic Press.
- Maddieson, I. (1997). Phonetic universals. In W. J. Hardcastle & J. Laver (Eds.), The Handbook of Phonetic Sciences, (pp. 619-639). Oxford: Blackwell Publishers.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word recognition. Cognition, 25, 71-102.
- Massaro, D. W. (1994). Psychological aspects of speech perception: Implications for research and theory. In M. A. Gernsbacher (Ed.), Handbook of psycholinguistics (pp. 219-263). San Diego, CA: Academic Press.
- 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.
- McClelland, A. G. R., & Pring, L. (1991). An investigation of cross-modality effects in implicit and explicit memory. Quarterly Journal of Experimental Psychology, 43A, 19-33.
- McClelland, J. L., & Elman, J. L. (1986). Interactive processes in speech recognition: The

- TRACE model. In J. L. McClelland & D. E. Rumelhart (Ed.), Parallel distributed processing: Explorations in the microstructure of cognition (pp. 58-121). Cambridge, MA: MIT Press.
- McQueen, J. M. (1998). Segmentation of continuous speech using phonotactics. Journal of Memory & Language, 39, 21-46.
- McQueen, J. M., Cutler, A., Briscoe, T., & Norris, D. (1995). Models of continuous speech recognition and the contents of the vocabulary. Language and Cognitive Processes, 10, 309-331.
- Miller, J. L., & Grosjean, F. (1997). Dialect effects in vowel perception: The role of temporal information in French. Language & Speech, 40, 277-288.
- Miller, J. L., & Volaitis, L. (1989). Effect of speaking rate on the perceptual structure of a phonetic category. Perception & Psychophysics, 46, 505-512.
- Morgan, J. L., & Saffran, J. R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. Child Development, 66, 911-936.
- Moscovitch, M., Winocur, G., & McLachlan, D. (1986). Memory as assessed by recognition and reading time in normal and memory-impaired people with Alzheimer's disease and other neurological disorders. Journal of Experimental Psychology: General, 115, 331-347.
- Mullennix, J. W., & Pisoni, D. B. (1990). Stimulus variability and processing dependencies in speech perception. Perception & Psychophysics, 47, 379-390.
- Musen, G., Shimamura, A. P., & Squire, L. R. (1990). Intact text-specific reading skill in amnesia. Journal of Experimental Psychology: Learning, Memory, & Cognition, 16, 1068-1076.
- Nakatani, L. H., & Dukes, K. D. (1977). Locus of segmental cues for word juncture. Journal of

- the Acoustical Society of America, 62, 714-719.
- Nakatani, L. H., & Schaffer, J. A. (1978). Hearing "words" without words: Prosodic cues for word perception. Journal of the Acoustical Society of America, 63, 234-245.
- Nusbaum, H. C., & Goodman, J. C. (1994). Learning to hear speech as spoken language. In J. C. Goodman & H. C. Nusbaum (Ed.), The development of speech perception (pp. 299-338). Cambridge, MA: MIT Press.
- Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words. Research on Speech Perception, Progress report no. 10. Speech Research Laboratory, Psychology Department, Indiana University, Bloomington, Indiana.
- Pierrehumbert, J. (1990). Phonological and phonetic representation. Journal of Phonetics, 18, 375-394.
- Pierrehumbert, J. (1994). Knowledge of variation. Paper presented at the Parasession on Variation, 30th meeting of the Chicago Linguistic Society, Chicago.
- Pilotti, M., Bergman, E. T., Gallo, D. A., Sommers, M., & Roediger, H. L. (2000). Direct comparisons of auditory implicit memory tests. Psychonomics Bulletin and Review, 7, 347-353.
- Pinker, S. (1989). Learnability and cognition: The acquisition of argument structure. Cambridge, MA: MIT Press.
- Pisoni, D. B. (1992). Some comments on invariance, variability, and perceptual normalization in speech perception. Proceedings of the 1992 International Conference on Spoken Language Processing. Banff, Canada.

- Poldrack, R. A., & Church, B. (1997). Auditory priming of new associations. Paper presented at the meeting of the Psychonomics Society.
- Poldrack, R. A., Selco, S. L., Field, J. E., & Cohen, N. J. (1999). The relationship between skill learning and repetition priming: Experimental and computational analyses. Journal of Experimental Psychology: Learning, Memory, & Cognition, 25, Jan 1999, 208-235.
- Quené, H. (1992). Durational cues for word segmentation in Dutch. Journal of Phonetics, 20, 331-350.
- Rueckl, J. G., & Olds, E. M. (1993). When pseudowords acquire meaning: Effect of semantic associations on pseudoword repetition priming. Journal of Experimental Psychology: Learning, Memory, & Cognition, 19, May 1993, 515-527.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). 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.
- Schacter, D. L., & Graf, P. (1986). Preserved learning in amnesic patients: Perspectives from research on direct priming. Journal of Clinical & Experimental Neuropsychology, 8, 727-743.
- Schacter, D. L., & Graf, P. (1989). Modality specificity of implicit memory for new associations. Journal of Experimental Psychology: Learning, Memory, & Cognition, 15, 3-12.
- Sheffert, S. M. (1998). Voice-specificity effects on auditory word priming. Memory & Cognition, 26, 591-598.
- Sloman, S. A., Hayman, C. A. G., Ohta, N., Law, J., & Tulving, E. (1988). Forgetting in primed

- fragment completion. Journal of Experimental Psychology: Learning, Memory, & Cognition, 14, 223-239.
- Sommers, M. S. (1999). Perceptual specificity and implicit auditory priming in older and younger adults. Journal of Experimental Psychology: Learning, Memory, & Cognition, 25, 1236-1255.
- Takeda, S., & Ichikawa, A. (1994). Analysis of prominence in spoken Japanese sentences and application to text-to-speech synthesis. Speech Communication, 14, 171-196.
- Vihman, M. (1996). Phonological development. Cambridge, MA: Blackwell Publishers.
- Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. Journal of Memory & Language, 40, 374-408.
- Woodward, A. L., & Markman, E. M. (1997). Early word learning. In W. Damon, D. Kuhn, & R. Siegler (Ed.), Handbook of Child Psychology, Vol. 2: Cognition, Perception, and Language (pp. 371-420). New York: Wiley.
- Woodward, A., Markman, E., & Fitzsimmons, C. (1994). Rapid word learning in 13- and 18-month-olds. Developmental Psychology, 30, 553-566.

## APPENDIX

Items for Experiment 1

bɪs	bæp
fauðz	fudz
gaɪb	gaʊr
ges	gæn
haɪz	hes
jitʃ	jev
jʌz	jig
kik	kæθ
lauθ	lɛl
lʌn	lɜːz
mɛk	maɪd
ʃeg	ʃauz
tɪrt	tem
tʃʌz	tʃɪdʒ
vɪrn	ves
wʌdʒ	wuθ

Items for Experiments 2, 3, and 4

<u>High-Low</u>		<u>Low-High</u>	
disgaɪb	tʌlgaɪb	dʌθlʌn	vezlʌn
dɪtlauθ	pemlauθ	dʒauðgæn	puθgæn
fʌɪkwʌdʒ	nælwʌdʒ	ðʌvkik	ʒɒbkik
nælfeg	fʌɪkfeg	gɪdʒvɜːn	kaɪðvɜːn
pemjitʃ	dɪtʃitʃ	ʒɒbmaɪd	ðʌvmaɪd
rɪzfaudʒ	sɜːkfaudʒ	kaɪðhes	gɪdʒhes
sɜːktʃɪʃ	rɪzʃɪʃ	puθtem	dʒauðtem
tʌljev	disjev	vezbæp	dʌθbæp
setʃwuθ	kebwuθ	ʃaɪpmɛk	dʒeθmɛk
sʌdjɪg	pɑːtʃɪg	ʃʌtʃhaɪs	gɜːghaɪs
sigfuːz	pɜːbfuːz	zɪθves	huːʃves
tʃɛnjʌz	pædʒjʌz	tautʃges	nɜːθges
keblɜːz	setʃlɜːz	dʒeθbɪs	ʃaɪpbɪs
pɜːbtʃʌz	sigʃʌz	gɜːgkæθ	ʃʌtʃkæθ
pɑːtʃauz	sʌdʃauz	nɜːθlɛl	tautʃlɛl
pædʒgaʊr	tʃɛngaur	huːʃtɜːt	zɪθtɜːt

## Footnotes

1. The experimenter stopped the task if the child gave no response or only irrelevant responses (e.g., "I don't know") for all of the first 8 of the 32 test items, or if the child declined to participate further, leaving 8 or more of the 32 items unheard. The rate of participant exclusion is similar to the rate found in previous research using this task (Church & Fisher, 1998).
2. Vocabulary data were missing for one child.
3. No vocabulary data were collected for these older children; the MacArthur Communicative Development Inventory used in Experiment 1 was designed for children up to 30 months old.
4. The corresponding item analysis differed slightly since each item had only one High and one Low phonotactic frequency syllable. We conducted a 2 (First vs. Second syllable) by 2 (HL vs. LH ordering of high and low frequency syllables) ANOVA, where Syllable was a within-items variable and HL/LH was a between-items variable. This analysis revealed an effect of syllable ( $F(1,62) = 12.78, p < .001$ ), and no effect of the position of the high frequency syllable ( $F(1,62) = 1.06, ns$ ) or interaction of syllable and the position of the high frequency syllable ( $F(1,62) < 1$ ).
5. As for Experiment 2, we conducted a corresponding item analysis: a 2 (First vs. Second syllable) by 2 (HL vs. LH ordering of high and low frequency syllables) ANOVA, where Syllable was a within-items variable and HL/LH was a between-items variable. This analysis revealed no effect of syllable ( $F(1,62) = 1.42, ns$ ) or of the position of the high frequency syllable ( $F(1,62) < 1$ ), but a significant interaction of these two factors ( $F(1,62) = 13.10, p < .001$ ), indicating that both first and second syllables were more accurately repeated if they were of high phonotactic frequency.
6. Poldrack and Church (1997) compared a semantic and a non-semantic study task in their

experiments. In the semantic study task, the subjects rated the likelihood that the two words in each pair would occur in the same sentence. In the non-semantic study task, subjects rated clarity of enunciation. The semantic study task created associative priming effects in all cases; the findings described in the text are from the non-semantic study task.

Table 1: Mean (SE) repetition accuracy for First and Second syllables, for Changed Context and New items, Experiment 2.

	Changed Context	New	Difference
First Syllable	.573 (.035)	.556 (.028)	.017 (.024)
Second Syllable	.518 (.030)	.455 (.029)	.064 (.021)

Table 2: Mean (SE) repetition accuracy for First and Second syllables, with Same or Changed contexts, Experiment 3.

	Same Context	Changed Context	Difference
First Syllable	.687 (.028)	.667 (.028)	.019 (.023)
Second Syllable	.677 (.025)	.626 (.025)	.050 (.016)

Table 3: Mean (SE) repetition accuracy for First and Second syllables, of High or Low phonotactic frequency, Experiment 3.

	High	Low	Difference
First Syllable	.708 (.024)	.645 (.028)	.063 (.015)
Second Syllable	.686 (.026)	.618 (.029)	.068 (.027)