Published in final edited form as:

Cognition. 2008 February ; 106(2): 833–870.

Influences of High and Low Variability on Infant Word Recognition

Leher Singh

Boston University

Abstract

Although infants begin to encode and track novel words in fluent speech by 7.5 months, their ability to recognize words is somewhat limited at this stage. In particular, when the surface form of a word is altered, by changing the gender or affective prosody of the speaker, infants begin to falter at spoken word recognition. Given that natural speech is replete with variability, only some of which is determines the meaning of a word, it remains unclear how infants might ever overcome the effects of surface variability without appealing to meaning. In the current set of experiments, consequences of high and low variability are examined in preverbal infants. The source of variability, vocal affect, is a common property of infant-directed speech with which young learners have to contend. Across a series of four experiments, infants' abilities to recognize repeated encounters of words, as well as to reject similar-sounding words, are investigated in the context of high and low affective variation. Results point to positive consequences of affective variation, both in creating generalizable memory representations for words, but also in establishing phonologically precise memories for words. Conversely, low variability appears to degrade word recognition on both fronts, compromising infants' abilities to generalize across different affective forms of a word and to detect similar-sounding items. Findings are discussed in the context of principles of categorization, both of a linguistic and non-linguistic variety, which may potentiate the early growth of a lexicon.

Indisputably, a crucial component of language acquisition involves learning the meanings of words. In the simplest terms, this refers to the process by which learners equate words that they hear with conceptual knowledge. When we consider the potential processes that enable word learning, it may serve us well to introspect on how we learn words in a new language as adults. As adults embarking on learning a second language, we approach the process by storing word-concept associations in memory. However, the fact that we have already successfully acquired one human language confers certain privileges that are unavailable to comparatively naïve infants learning their first language. First, as adults, we can avail of certain assumptions about human languages in general. Specifically, we know that human languages are divisible into words, clauses and phrases, even though speech arrives at our perceptual receptors as a continuous stream of sound. This knowledge compels us to mentally punctuate the signal and generate smaller-sized units that may be easier to parse. Second, we know something about the purpose served by words in a language. Specifically, we know that a word is a subclausal unit that maps onto a concept in our native language, arguably forming the most fundamental unit of the language code employed to impart meaning. Finally, the equation of sound and meaning is made easier for us as we are explicitly aware of the ways in which a particular language expresses meaning, i.e. how changes in sound determine changes in meaning in the language we are attempting to learn. For example, if we opt to learn a tonal language, we know

Address for correspondence: Leher Singh 635 Commonwealth Avenue Department of Communication Disorders Boston University
Boston MA 02215 (617) 343-7491 leher@bu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

that we need to attend to certain types of pitch change while defining words. This knowledge allows us to pay close attention to changes in sound that affect meaning and less attention to those changes that are not meaningful. Therefore, at a very general level, we possess a few basic definitions of a word before even approaching the task of word learning in a second language. This affords us a comfortable degree of predictability about the status and possible form of words in the language we are trying to learn at the outset. By contrast, infants learning their first language lack knowledge of how and why words are relevant to language use. They do not know *a priori* that human language can and should be parsed into words nor do they know how their particular language chooses to alter the form of a word in order to alter its meaning. Therefore, prior to mapping sound to meaning, young infants must derive certain facts about human language as well as about their own particular language before they can 'discover' words in the speech stream.

The process of word discovery in infancy is complicated by two widely documented problems. First, infants must resolve the segmentation problem, whereby they have to separate large-scale units into individual components that correspond to words. This poses a problem because speech, unlike written language, does not unravel with convenient pauses inserted between words. Therefore, infants have to establish the location of word boundaries prior to knowing the meaning of most words. Given that parents produce a negligible percentage of total words in isolation to their children and that vast majority of their speech consists of multiword utterances (Van der Weijer, 1998), the segmentation problem is one that infants must overcome without assistance from their caregivers. Second, infants must contend with inordinate variability in speech, caused by a 'lack of invariance' (Klatt, 1989). It is well known by scholars of human speech perception and architects of automatic speech recognition devices that human speech is a calculus of cues that interacts in notoriously complex and unpredictable ways with human language. While we all have a reliable and stable store of linguistic knowledge that we consult when we produce or perceive speech, the way in which we produce or perceive linguistic units varies in unreliable ways both between and within individual speakers. Within speakers, depending on factors such as a speakers' affective state or distance from the listener, linguistic units will sound physically different across varying contexts. Across speakers, different vocal tract characteristics lead to the units of speech being realized in physically distinct ways based on factors such as gender and voice quality. Even if we consider the smallest units of speech, a phonetic segment can change its form based on surrounding segments as a result of assimilation processes (Gaskell & Marslen-Wilson, 1996; Gow, 2001; 2002). As a result, it is incredibly difficult to capture invariant linguistic units amidst the mire of human speech.

Both the segmentation and variability problems conspire to produce a challenge that would seem intractable for young minds. Amidst profound variability and its ensuing indeterminacy, how do infants establish the way in which words should be represented in memory? Painting in broad strokes, this problem amounts to a task of reducing the dimensionality of the input to its linguistically relevant (phonemic) dimensions and categorizing incoming sounds accordingly. However, in order to categorize speech according to dimensions that are phonemic, infants must first know which dimensions of sound are phonemic in their native language. One potential way to arrive at this knowledge is to appeal to meaning: Changes in sound that necessarily accompany changes in meaning are linguistically relevant. However, paradoxically, infants learn to recognize words in fluent speech before they map these words onto meaning with any regularity. In a now seminal study by Jusczyk & Aslin (1995), it was revealed that infants learn to track, encode and recognize repetitions of novel words, with which meaning has not yet been associated, by 7.5 months. In this study, infants were familiarized with repetitions of two words using the Headturn Preference Procedure. They were then exposed to sentences containing those words interspersed with an unfamiliar set of sentences. Infants listened longer to sentences that contained the familiarized words than to unfamiliar

sentences, providing the first empirical evidence of speech segmentation and word recognition in infancy. Therefore, months before they are able to understand the meanings of words, infants continuously archive memories of words they hear, perhaps heralding the true point of origin of lexical development.

Even though infants develop word knowledge as early as 7.5 months, there is considerable fragility in their word recognition abilities. Specifically, 7.5-month-old infants appear to encode words in fine phonetic and acoustic detail, which compromises their ability to detect novel instances that are acoustically dissimilar. Consequently, at this stage, infants do not recognize a word spoken by a female if they were trained on an instance of the word spoken by a male (Houston & Jusczyk, 2000) nor do they recognize a word spoken with positive affect if they were trained on the word in neutral affect and *vice versa* (Singh, Morgan & White, 2004). In a study by Singh et al. (2004), 7.5 month old infants were familiarized with two words, one in a happy tone of voice and another in a neutral tone of voice. Infants were then tested on their recognition of both words in the context of sentences. As in Jusczyk & Aslin's (1995) study, during the recognition phase, some of the sentences contained the familiarized words and some did not. The difference in infants' listening times to the two types of passages yielded an index of word recognition. One distinction between the design of this experiment and that of Jusczyk & Aslin (1995) was that during the recognition phase, half of the infants heard all passages (groups of six sentences) in happy affect and half of the infants heard all the passages in neutral affect. Results demonstrated that at 7.5 months, infants only recognized happy familiarization words in happy passages and neutral familiarization words in neutral passages, revealing a matching effect. Essentially, infants failed to recognize the same word when it differed in affect between the two phases of the experiment. Over the succeeding three months, infants' ability to recognize dissimilar encounters of the same word appear to markedly improve. At 10.5 months, infants succeed in recognizing instances of a word that were mismatched in vocal affect (Singh et al., 2004) or speaker gender (Houston & Jusczyk, 2000) indicating the point at which infants could learn to appropriately generalize to novel, dissimilar tokens of the same word in a linguistically mature fashion (Houston & Jusczyk, 2000; Singh et al., 2004).¹

The transition from 7.5 to 10.5 months marks a concurrent development in phonetic perception where infants graduate from a universalist to a language-specific frame of reference. While infants begin life with a perceptual apparatus designed to detect phonetic contrast regardless of whether it proves to be phonemic in the native language, this apparatus is pruned over the second half of infancy to selectively appreciate phonetic changes that are phonemic in the native language (Best, 1995; Kuhl, 1996; Werker & Tees, 1984; 1999) at the cost of attending to those contrasts which are not. A rich and detailed literature, inspired by an original study by Werker and Tees (1984) has allowed us to chronicle the development and elaboration of infants' phonological store over the first year of life. This literature has uncovered great linguistic strides made by early learners in the establishment of a native phonological repertoire, which is undoubtedly a necessary prerequisite to acquiring other formal aspects of language. While infants develop a language-specific orientation towards the end of their first year, the exact time course of this transition interacts with a number of other factors, such as the statistical frequency of the phonemes being tested (Anderson, Morgan & White, 2003; Maye, Werker & Gerken, 2002), the relationship of the native language to the contrasts being tested (Best,

¹An interesting addendum to have emerged from these studies is that the way in which infants expressed word recognition accorded with their perceptual preferences for different affective styles. In word recognition tasks when affect was varied, infants consistently and reliably expressed recognition of words familiarized in a happy voice by attending significantly *longer* to passages containing that word relative to the baseline. By contrast, infants expressed recognition of words familiarized in neutral affect by attending for a significantly *shorter* time to passages containing the familiarized word relative to baseline. Both of these directions of preference are considered to be valid indices of word recognition, as inhibition of attention to a stimulus necessitates first recognizing the stimulus as one was previously encountered in a style judged to be unappealing.

1995), and the particular training conditions under which infants are familiarized with phonemes (Maye, et al, 2002).

While it remains unclear exactly how infants learn which contrasts are phonemic, one potential causal mechanism by which infants may establish phonemic boundaries is the statistical frequency with which particular segments occur in the input. In a study designed to assess infants' sensitivities to such frequencies, Maye, et al. (2002) demonstrated an impressive capacity on the part of 8 month old infants to draw phonemic distinctions by capitalizing on distributional cues in the input. In this study, when infants were exposed to phonetic continua that assumed a bimodal distribution, infants formed two phonetic categories; when exposed to phonetic continua assuming a unimodal distribution, they formed a single phonetic category. This provides a compelling causal mechanism by which infants may develop native phonetic categories amidst considerable natural variability in the production of phonemes in the linguistic environment. Therefore, it remains clear that infants undergo crucial developments between 6 and 12 months, during which they appear to develop a native perceptual filter through which incoming phonetic segments are classified. This process may be guided by the distributional profile of particular phonetic segments in the input, revealing the valuable contributions of early plasticity to the discovery of native phonological organization.

This emergence and refinement of native phonetic categories between 6 and 12 months seems to coincide in part with the development of robust word recognition skills between 7.5 and 10.5 months. Therefore, infants, by the end of their first year, have learned to ascribe relevance to the acoustic properties of phonemes and to the acoustic properties of words based on the underlying phonological organization of their language. A primary goal of the current set of studies is to determine what factors may facilitate or impede the process of ascribing relevance to the properties of words and in particular, how the natural variability encountered in typically infant-directed speech may guide such a progression.

Even though infants develop the ability to recognize familiar words in fluent speech at 7.5 months (Jusczyk & Aslin, 1995), they have a strong propensity to retain talker- or context-specific details in memory to their own detriment in these tasks. This tendency appears to result in narrow lexical categories, in which words are defined in terms of both phonetic and acoustic characteristics. Consequently, infants show matching effects and successfully detect only familiarized words that are both phonetically and acoustically similar. Such matching effects are evident regardless of whether similarity is realized by complex constructs such as talker gender and vocal affect, which are generally characterized by a constellation of spectral and temporal changes, or by simpler lexically irrelevant dimensions such as absolute pitch (Singh, White & Morgan, *in press*). Therefore, at early stages of word recognition, infants appear to fuse phonological and perceptual characteristics, storing in memory highly specific composites. In theory, it would behoove learners to adopt such a conservative approach to early word learning as languages differ in the kinds of acoustic cues they exploit to communicate meaning. Therefore, by design, infants cannot arrive with a prescribed set of rules mandating which acoustic details to consider and which to disregard in structuring a native lexicon. As a consequence, they may cautiously encode all surface details in the event that they may prove lexically relevant. Later, as infants gain more exposure to words in ever increasingly varying forms, they may learn which dimensions of sound co-vary with meaning and which dimensions vary orthogonally to meaning.

In other words, the transition from fragile to robust word recognition observed between 7.5 and 10.5 months may reflect infants' mastery of the interaction of acoustic-phonetic cues and meaning within a given language. However, implicit in this account is the notion that infants view words as lexical items at this later stage. In light of the fact that infants' word recognition skills appear to be robust at 10.5 months, when they may possess only a very modest

comprehension vocabulary (Benedict, 1979), it seems unlikely that knowledge of word meanings, in and of itself, strengthens word recognition at this age. It is possible that older infants may develop mature word recognition skills without appealing to meaning at all. Instead, they may succeed at the task based on the type of experience they have accrued with words. Specifically, the diversity of experience they have had with words may assist them in generalizing appropriately across encounters of the same word. Accordingly, they may capitalize on variability along lexically irrelevant dimensions to identify the invariants, which are likely to be germane to lexical identity.

If it is true that older infants exploit variability in the speech stream to learn how to generalize across encounters of words obviating any need to appeal to meaning, why do young infants perform poorly at word recognition in the face of surface variability? One possible reason is that they have simply had less experience with words, and therefore, less opportunity to observe the infinitely varying forms that words can assume. Therefore, given a high degree of inexperience and uncertainty, young infants may take their cues from the conditions of the task. They may assume that all covarying properties are relevant to categorizing the word. Therefore, when they hear a single word repeated in a particular affect, they may assume that both the phonological and affective properties of the word contribute equivalently to its identity. Over the next few months, as they experience a greater diversity of word forms, they may identify the invariant properties of words for their language and focus only on those dimensions of sound when categorizing words. This account circumvents the role of meaning in the maturation of spoken word recognition, and therefore might help to explain how infants possess relatively mature word recognition skills at 10.5 months in the absence of a substantial vocabulary. By this account, just as limited experience with the same word (e.g. only hearing the word in a single type of affect) may lead to narrow categories for that word, it is equally likely that diverse experience with a word (e.g. hearing a word spoken in many different types of affect) may lead to broad categories for that word.

Just as we observe infants' tendencies to over-rely on surface form in early lexical categorization, they show similar tendencies in early cognitive categorization. Analogous studies on early object categorization reveal that infants initially attend to all surface form details, in the event that they may be meaningful (Oakes & Madole, 2003). Later, they learn to disregard those regularities that emerge adventitiously, and only consider those that are reliably co-occur (Madole & Cohen, 1995, Needham, Dueker & Modi, 2002). This potential bias to attach relevance to details that reliably co-occur across encounters allows infants to progress from attending to all feature correlations to only those that are consistently supported by their interaction with the world (Madole & Oakes, 1999). Therefore, nascent categories are highly vulnerable to effects of surface form variation across exemplars because it is assumed that similarity in form across encounters is relevant.

Therefore, while the matching effects observed in word recognition tasks completed by younger infants may appear linguistically immature, they may simply reflect immature categories, rather than immature processes. The principles that infants may employ to form these categories may be relatively mature, in the sense that they are based on structural correlations in the input. When the opportunity to derive those correlations is limited, either by having had few encounters with a word and/or having a perceptually narrow range of encounters, infants may rely heavily on structural details reinforced in a particular task. Similarly, the actual categories that infants demonstrate evidence of may be task- and age-dependent, whereas the means by which infants construct categories may remain constant over time. By this account, categories would automatically stabilize as infants cull increasing amounts of information about which characteristics covary reliably and which do not. This invites an overarching question, which underlies one of the most clamorous debates in early

cognitive and linguistic development: When forming categories, whether cognitive or linguistic (or both), does the structure reside in the input or the infant?

To train this question towards the challenges of early word recognition, it is possible that in previous studies, by familiarizing infants with a word in a single affect (either happy or neutral) or spoken by a single gender (Houston & Jusczyk, 2000; Singh et al., 2004), this may have reduced the scope of any lexical category that infants may have formed during familiarization. Instead, if infants were familiarized with each word in a variety of affects for example, would they have succeeded in forming more generalizable memory representations for words? If infants' emergent lexical categories are in fact dynamic and responsive to experience, one might expect them to update their representations based on the kind of information they receive during the task. Increased variability across instances of a word may indicate to infants that highly varying dimensions are irrelevant to lexical identity and may lead to the formation of more robust representations. This hypothesis is investigated in the following study to determine the extent to which infants' early lexical categories are malleable (yet opportunistically so) early in development.

Intuitively, one might expect increased stimulus variability to hinder categorization in infants by increasing processing load during the task. Given that their ability to categorize words at 7.5 months is already fragile, this increase in load may serve only to usurp valuable resources and disrupt word recognition. However, the current hypothesis is predicated on the notion that the observed fragility in young infants' word recognition abilities is in part ascribable to the fact that their lexical categories are in their inception. The most primitive categories in particular may benefit greatly from increased variability in surface form because naïve learners may turn to cues provided by a task in their search for defining characteristics more than experienced learners

Furthermore, if we consider the privileges and constraints attached to high variability within the context of input to infants, infant-directed speech is replete with affective variation (Fernald, 1993; Trainor, Austin & Desjardins, 2000), suggesting that affectively variable speech may more closely approximate the kind of language infants actually hear than affectively uniform speech. It is highly improbable that children are ever introduced to instances of words only in a single affect, given the characteristic modulations that adults often incorporate into their speech to infants. Therefore, familiarization with a wide range of tokens more realistically simulates the prosodic undulations of a natural mother-infant dyad.

In Experiment 1, 7.5 month old infants were tested on their ability to recognize words in fluent speech amidst high variability in a procedure similar to that of Jusczyk & Aslin (1995) and Singh, et al. (2004). Infants were familiarized with one word in variable affect, another word in constant affect (either positive or neutral) and then tested on their recognition of these words in fluent speech. The recognition stimuli were spoken in constant affect (either positive or neutral), which served as a between-subjects variable. However, the affect of the recognition passages was always *mismatched* with the affect of the constant word. For example, a particular infant might hear one word in variable affect, one word in positive affect during familiarization and then all recognition passages in neutral affect. Based on the results of Singh et al. (2004), it was expected that the infants would recognize the word they heard in variable affect, owing to the hypothesized advantages of variable affect, yet that they would fail to recognize the word in constant affect as it was mismatched across familiarization and recognition.

Experiment 1

The goal of the present study is to determine whether infant word recognition is aided by the presence of increased variability in surface form. To investigate this, infants were familiarized with one word spoken in several different emotions (variable word) and another word in a

single emotion (constant word). They then heard each word in a series of passages interspersed with unfamiliar passages (sentences containing words to which the infant was not familiarized). In this design, the affect of the passages (happy or neutral) always mismatched the affect of the constant word (happy or neutral). From previous studies, it has been established that infants at this age do not recognize words across encounters that contrast in surface form as a result of changes in factors such as affect, talker gender, and pitch (Houston and Jusczyk, 2000; Singh, et al., 2004, Singh, et al., under review). It was predicted that infants would recognize the word familiarized in variable affect in both types of passages. Furthermore, it was predicted that infants would fail to recognize the other word, given that it was mismatched in affect across the familiarization and recognition phases of the experiment.

Participants

Forty full-term, English-exposed 7.5 month olds participated in the study (22 males and 18 females), recruited from Rhode Island Health Department records. The mean age of participants was 231 days (range = 224 days to 248 days). Eight additional infants were tested and data were discarded because of inattention or technical difficulties.

Stimuli

Stimuli consisted of four monosyllabic words (“bike”, “hat”, “tree” and “pear”) and four passages, each containing six sentences. The test passages and the constant affect familiarization words were identical stimuli to the stimuli used in previous investigations (Singh et al., 2004). To produce the variable affect words, the speaker was asked to produce each word in happy, neutral, sad, angry and fearful affect. Five repetitions in each affect were produced, generating a total of 25 tokens per word. The three most clear and demonstrative tokens for each emotion were selected and formed the stimulus set of 15 variable affect words. All stimuli were produced by a trained theater actor so that the intended emotions were convincingly conveyed. In addition, the speaker was addressing her own infant while producing the stimuli to ensure that they were spoken in an infant-directed register. The stimuli for all studies were produced by the same speaker as in previous studies (Singh, et al., 2004). In addition, the test passages for every study were identical to those used in previous studies (Singh, et al., 2004).

Acoustic analyses were conducted on both the variable and single affect familiarization stimuli and are graphed in Figures 1a and 1b based on the most significant communicators of vocal affect: mean fundamental frequency and fundamental frequency range, which are particularly instrumental in distinguishing positive affect from neutral affect (Banse & Scherer, 1996; Scherer, 1986; Williams & Stevens, 1972). Mean fundamental frequency, fundamental frequency range, maxima and minima are presented in Table 1. As can be seen in Figures 1a and b, the variable affect words form an acoustically heterogeneous set relative to the constant affect words for all measures. Figure 1a shows the mean fundamental frequency for each familiarization word for variable, happy and neutral familiarization items and each point represents an individual familiarization token. As Figure 1a shows, the mean fundamental frequencies of the variable tokens are much broadly distributed than those of each of the happy and neutral tokens. In fact, the range of values spanned by the variable tokens encompasses that spanned by both the happy and neutral group combined. A comparison of variances using Levene's test for Equality of Variances revealed that the variances of each group of affective words were significantly different from that of the variable words, $F(1, 118) = 55.7, p < .0001$ (happy and variable) and $F(1, 118) = 154.1, p < .0001$ (neutral and variable). Similar results were found for fundamental frequency range (displayed in Figure 1b), $F(1, 118) = 8.9, p < .01$ (happy and variable) and $F(1, 118) = 68.45, p < .0001$ (neutral and variable). These findings corroborate the patterns depicted in Figures 1a and 1b, that along dimensions related to fundamental frequency, the variable words constituted a broader set of tokens than either of the single affect

sets of words. All familiarization stimuli were equated for mean amplitude, duration and vowel length.

Comparisons of the happy and neutral tokens revealed that minimum and maximum F_0 were higher in happy words relative to neutral words, $t(59)=9.39$, $p<.0001$ and $t(59)=33.58$, $p<.0001$ respectively. Mean F_0 was higher in happy words than in neutral words, $t(59)=29.55$, $p<.0001$. Pitch range of happy words also exceeded that of neutral words, $t(59)=15.18$, $p<.0001$.

For sentences (see Table 2), happy tokens embedded in sentences had higher F_0 minima and maxima compared with neutral tokens, $t(23)=8.96$, $p<.0001$ and $t(23)=25.93$, $p<.0001$ respectively. In addition, pitch was higher and more variable in happy sentences relative to neutral sentences, as indexed by F_0 mean and range, $t(23)=25.47$, $p<.0001$ and $t(23)=21.88$, $p<.0001$ respectively. Finally, analyses revealed a higher proportion of high-frequency energy in happy sentences than in neutral sentences, $t(23)=7.40$, $p<.0001$, consistent with previous acoustic profiles of happy and neutral speech at the level of the phrase (Banse & Scherer, 1996; Scherer, 1986). The duration of happy sentences did not differ significantly from the duration of neutral sentences, indicating that speech rate did not differ significantly across affect. In addition, the durations of the target words within the carrier sentences did not differ between happy and neutral passages.

Apparatus

Testing was conducted in a three-walled testing booth within a sound-treated testing room. Each wall of the booth was 120 cm wide. A chair was positioned at the open end of the booth where the parent sat with the infant on his/her lap. The infant sat approximately 110 cm from the front of the booth. Yamaha bi-amplified loudspeakers were located behind both side walls of the booth. At the infants' eye level, 86 cm above the floor, a white light was mounted on the front wall. Each of the side walls had a similar blue light at the same level. A Panasonic CCTV video camera (model WV-BP330) was mounted behind the testing booth 12.3 cm above the yellow light. In a separate control room, a Panasonic monitor (WV-5410) was connected to the video camera in the testing booth. The participants were displayed on the monitor in the control room, where the experimenter judged infants' looking, pressing buttons on the mouse of a Windows computer to control the customized experimental software. The computer was equipped with a Sound-Blaster compatible soundboard connected to the amplified speakers. Speech stimuli were set at conversation level (75 dB) using a Digital sound level meter.

Procedure

Infants were tested using the Headturn Preference Procedure (HPP) (Kemler Nelson, Jusczyk, Mandel, Myers, Turk, & Gerken, 1995). The infant was seated on the parent's lap facing the center light. The parent listened to instrumental music over Panasonic headphones to mask the stimuli. Each trial began with the center light flashing until the experimenter judged that the infant fixated on the flashing light. At that point, this light was turned off and one of the side lights began to flash to attract the infant's attention to the side. Side of presentation was randomized across trials, so that all stimuli occurred on both sides. After the infant turned to look at the flashing side light, the speech stimuli for that trial began to play. The sound continued to play and the side light remained on for the duration of the infant's fixation on the light. Each trial continued until the infant looked away for two seconds, or until 20 seconds of looking time had been accumulated during that trial. If the infant looked away, but then looked back within two seconds, the trial continued. If the infant's looking time was below 2 seconds, the trial was repeated with a new randomization of the trial stimuli; otherwise, the procedure advanced to the next trial.

Familiarization began with trials alternating between the two target words. Once the infant had exceeded 30 seconds of looking time with one word, all subsequent familiarization trials presented the alternate word. This modification of the HPP was instituted to ensure that differences in looking times during recognition testing could not be due to different amounts of familiarization with the two target words. When the infant reached 30 seconds of looking time with the second word, the test phase began.

Recognition testing consisted of four blocks of trials, each block containing one trial with each of the four passages. The order of passages within each block was randomized for each infant. In addition, the order of sentences within passages was also randomized on each trial. The test procedure was similar to the familiarization procedure, except that the side light continued to flash while infants were fixated on the light. As in the familiarization phase, if the infant continued to look at the light for 20 seconds, the trial ended automatically and the next trial began. Similarly, if the infant failed to look at the side light for at least 2 seconds, the trial was automatically repeated. A minimum criterion of 2 seconds was necessary to allow the infant to hear at least one token of the target word in a sentence.

Target words for each infant were either “bike” and “hat” or “tree” and “pear.” The within-subjects manipulation was the affect of the target words: all infants heard one target word in variable affect and the other in either happy or neutral affect. The affect (variable/constant) of the target word and the particular pair of words infants were familiarized with were counterbalanced across subjects. In the interests of clarity, a summary of the factors manipulated in this and subsequent experiments is displayed in Table 4.

During the test phase, infants heard four passages. Half the infants heard all passages in happy affect and half heard all passages in neutral affect. However, the affect of passages was always mismatched with the affect of the constant familiarization word. For example, an infant who heard “bike” in variable affect and “hat” in neutral affect heard all recognition passages in happy affect. Therefore, there was one between-subjects condition, passage affect during recognition (happy or neutral), and one within-subject condition, word affect during familiarization (variable or constant).

Results and Discussion

To calculate infants' recognition of familiarized words, recognition scores were computed, which were calculated by subtracting infants looking times to passages containing unfamiliar words from looking times to passages containing familiarized words. Evidence of word recognition was inferred from a recognition score that departed significantly from zero. Although analyses were conducted on recognition score, raw data are available in Table 5 for each experiment and condition therein. Recognition scores with standard error values are plotted in Figure 2 for happy passages and Figure 3 for neutral passages. For infants who were familiarized with words in neutral and variable affect and tested on recognition of those words in happy passages (see Figure 2) there was a main effect of word affect, $F(2,38)=10.55, p<.0001$. In this condition, infants listened longer to passages containing words they originally heard in variable affect than to passages containing unfamiliar words, $F(1,19)=4.31, p<.05$. Eleven of twenty infants looked longer at sentences containing words familiarized in variable affect relative to unfamiliar passages. However, in this condition, infants also showed reduced looking times for words familiarized in neutral affect, even though they were embedded in happy test passages $F(1,19)=6.78, p<.05$. Twelve of twenty infants showed this pattern of results. This tendency on the part of infants to inhibit their attention to stimuli presented in neutral affect was observed reliably in pilot testing and in previous studies investigating recognition of words spoken in neutral affect in both 7.5 and 10.5 month old infants (Singh et al., 2004). Given the reliability with which this reverse pattern of looking time has been observed with stimuli spoken in neutral affect, it is hypothesized that infants inhibit their

attention to neutral affect stimuli because infants typically disprefer neutral and negative affect stimuli in speech perception tasks (Fernald, 1993; Kitamura & Burnham, 1998; Singh, Morgan & Best, 2000). By contrast, infants listen selectively to speech spoken in positive affect, perhaps accounting for increased looking times relative to baseline for familiarization items spoken in positive affect. Both reduced and increased looking times relative to baseline are indications of word recognition, but again signify that infants express their recognition of a word in ways consonant with their listening preferences.

In the neutral passage condition, infants were exposed to one target word in happy affect and the other in variable affect. They were then tested on recognition of these words by listening to a neutral set of passages containing both words as well as unfamiliar passages. In this condition, infants looked longer at passages containing words familiarized in variable affect relative to unfamiliar words, $F(1,19)=7.23$, $p<.05$ (see Figure 3). Eleven of twenty infants showed this pattern. However, in this condition, infants also showed increased looking times for words familiarized in happy affect, even though they were embedded in neutral test passages, $F(1,19)=8.65$, $p=.01$. Fourteen of twenty infants showed this pattern of results.

In summary, infants were able to recognize words familiarized in variable affect in the context of both happy and neutral test passages, suggesting that high variability in familiarization assists infant word recognition. However, the most surprising result of this experiment was that unlike in previous investigations, 7.5-month-old infants recognized words presented in a single affect during familiarization. This was inconsistent with earlier predictions because these single affect tokens were always embedded in passages that contrasted with the affect of familiarization stimuli. Therefore, the effects of high variability in one word appeared to extend to the other word presented in the session, even though that word was produced with minimal variability. The finding that the benefits of a high variability set propagated to a low variability set implies that the focus of infants' attention may have shifted from exemplar-specific details to category-based details for both words in the test session. In other words, infants appeared to be able to disregard exemplar-specific surface details in favor of category-level representation, resulting in the appearance of relatively mature word recognition skills.

Infants appear to show relatively abstract word recognition skills in this experiment, suggesting that while they are known to retain considerable episodic detail about words, they are able to focus on phonological characteristics of both words presented in the same experimental session under particular conditions. In light of the finding that high variability across tokens facilitates subsequent word recognition, these results raise the issue of whether low acoustic variability during familiarization degrades performance on word recognition tasks. Previously, it has been reported that low variability during familiarization leads infants to over-emphasize perceptual similarity along dimensions that are lexically irrelevant (Houston, 1999; Singh et al, 2004). To what extent does this over-reliance on surface detail compromise spoken word recognition? One possible cost of this is that infants may over-emphasize affective dimensions at the expense of attending to more subtle cues such as fine phonetic detail. The following experiment assesses the effects of low acoustic variability on infants' detection of phonological equivalence by seeking evidence of false recognition of similar sounding words that are perceptually confusable (i.e. produced with the same affective prosody) yet lexically distinct (i.e. minimal pairs).

Previous investigations of false recognition in 7.5-month-old infants have revealed that they are highly sensitive to phonetic detail in spoken word recognition tasks and that they do not confuse minimal pairs. These studies showed that infants do not equate phonetic variants that differ by the onset phoneme, e.g. they do not perceive "tup" to be an instance of "cup" (Jusczyk & Aslin, 1995). Similarly, they are sensitive to final consonant quality, which is typically less perceptible, and do not show false recognition of variants that differ in the final consonant

either (Tincoff & Jusczyk, 1996). However, in these studies conducted by Aslin, Jusczyk & Tincoff, the stimuli were spoken in typical infant-directed speech and therefore each familiarization set consisted of acoustically variable instances of a word, and were possibly more similar to the variable stimulus set used in Experiment 1. The high variability inherent in the familiarization sets may have assisted infants in attending to invariant phonological details and disregarding lexically irrelevant details, leading to a similarly high level of performance to Experiment 1. The current focus of investigation is in the context of low stimulus variability, whether infants over-represent surface details that reliably co-occur with phonological structure and in particular, whether this over-reliance leaves in its wake a neglect of more subtle phonemic detail. In the following experiment, infants are introduced to a phonetic variant of each target word produced in a single affect (e.g. “dike” consistently produced with happy affect and “gat” consistently produced with neutral affect). They were then tested on recognition of target words (e.g. “bike” and “hat”) in sentences spoken in either happy or neutral affect. Therefore, no word was introduced with high stimulus variability. This design was very similar to that of Singh, et al. (2004) except that here, infants were familiarized with variants of the target word. However, the design was also similar to Experiment 3 in Jusczyk & Aslin (1995) where infants were familiarized with variants of the words on which they were later tested, although unlike in that experiment, in the present experiment the variability with which words were introduced was varied between words rather than within a word.

Experiment 2

The purpose of the following experiment was to investigate whether false recognition might be observed in 7.5-month-old infants when the recurrence of surface properties of the familiarization tokens directs their attention to non-phonemic surface properties, such as affect. If infants show evidence of false recognition, this would again suggest that they are uncertain about the determinants of lexical relevance at this age, and are influenced by nonphonemic features at the expense of phonemic features. In this case, it would appear that low stimulus variability in familiarization not only results in increased matching based on nonphonemic similarities but also that it actively compromises matching based on phonemic similarities.

Participants

Thirty-two full-term, English-exposed 7.5 month olds participated in the study (13 males and 19 females), recruited from Rhode Island Health Department records. The mean age of participants was 234 days (range = 224 days to 249 days). Six additional infants were tested and data were discarded because of inattention.

Stimuli

Stimuli consisted of four monosyllabic non-words (“dike”, “gat”, “bree” and “gare”). These non-words were designed to match the four target words used in a previous study (Singh, et al., 2004), “bike”, “hat”, “tree”, and “pear”. The process of generating non-words was identical to that used to generate the words used in the previous study, where the same speaker was asked to simulate happy and neutral affect. During the recording session, the speaker was provided with the happy and neutral target words used in the previous study and asked to mimic the affective prosody of the target words, resulting in as close a match as possible between the target words used in previous studies and the non-words used in the current study. Acoustic analyses for all happy and neutral words are presented in Table 3. In addition, there were four passages, each containing six sentences. The test passages were the identical to those used in Experiment 1 and consisted of passages containing the words ‘bike’, ‘hat’, ‘tree’, and ‘pear’.

The non-words were then acoustically analyzed. As described previously, two of the primary acoustic correlates of vocal affect are mean F_0 and F_0 range (Banse & Scherer, 1996). Each of these was measured for happy and neutral non-words. Acoustic measures for all words are shown in Table 2. For individual words, mean F_0 was higher in happy words than in neutral words, $t(59)=-9.44$, $p<.0001$. Pitch range of happy words also exceeded that of neutral words, $t(59)=-7.99$, $p<.0001$. However, there was no difference in the relative durations or intensities of happy and neutral words.

Apparatus and Procedure

Familiarization words for each infant were either “dike” and “gat” or “bree” and “gare.” The within-subjects manipulation was the affect of the familiarization words: all infants heard one familiarization word in happy affect and the other in neutral affect. The assignment of affect of each familiarization word and the particular pair of words infants were familiarized with were counterbalanced across subjects.

During the test phase, infants heard four passages. Half of the infants heard all passages in happy affect and the other half heard all passages in neutral affect. Therefore, there was one between-subjects condition, passage affect during recognition (happy or neutral), and one within-subject condition, word affect during familiarization. The procedure was otherwise identical to that of Experiment 1.

Results and Discussion

As before, difference scores were used as an index of word recognition although raw data are available in Table 5. In addition, within-group analyses were performed individually for infants who heard happy and neutral recognition passages. Infants who heard all passages in happy affect showed a significant main effect of word affect, $F(2,30)=4.6$, $p<.05$ (see Figure 4). These infants showed increased looking times for passages containing words familiarized in happy affect relative to unfamiliar passages, $F(1,15)=5.9$, $p<.05$. Eleven of sixteen infants showed this type of preference. There was no significant difference in looking times to passages containing words familiarized in neutral affect relative to those containing unfamiliar words, $F(1,15)=.06$, NS. Therefore, infants falsely recognized happy non-words as exemplars of happy targets but did not falsely recognize neutral non-words as exemplars of happy targets. This pattern of results is highly similar to those of previous studies (Singh, et al., 2004) when infants were familiarized with the actual targets, rather than a similar-sounding variant.

Infants who heard all passages in neutral affect showed a marginally significant effect of word affect, $F(2,30)=3.19$, $p=.055$. Individual planned comparisons revealed significantly reduced looking times for passages containing words familiarized in neutral affect relative to unfamiliar passages, $F(1,15)=6.87$, $p<.05$ (see Figure 5). Thirteen of sixteen infants showed this pattern of results. They showed no significant difference in looking times to neutral passages containing words that corresponded to happy non-words presented in familiarization, $F(1,15)=.61$, NS. Therefore, infants showed false recognition of neutral targets that corresponded to neutral non-words in familiarization, but showed no recognition of neutral targets that corresponded to happy non-words in familiarization.

These findings reveal effects of similarity-based matching that closely resemble those observed in previous investigations of spoken word recognition. However, in this experiment, infants falsely recognized non-words that were affectively matched but phonemically distinct from their targets. Therefore, it appears that when the structure of the familiarization set draws infants' attention to certain dimensions of similarity, their ability to detect more subtle properties such as phonological equivalence may suffer. At first blush, evidence of false recognition at this age using the current procedure may appear inconsistent with findings from

Jusczyk & Aslin (1995) and Tincoff & Jusczyk (1996) that reveal a high degree of precision in infants' encoding and retrieval of phonological detail. In a similar task, infants in their studies failed to falsely recognize items that departed from the target items by one or more phonological features, either in the onset or final consonant. However, when evaluated in the context of Experiment 1, it appears that the high surface variability with which false phonetic variants were introduced in Jusczyk & Aslin's (1995) procedure may have contributed to more robust performance on the part of infants. Just as high stimulus variability may have enabled more robust word recognition in their study, low stimulus variability may have compromised encoding and retrieval of words presented during familiarization in the present experiment. Therefore, these seemingly discrepant findings may in fact fit compatibly into the same fold, perhaps revealing a complementarity of effects caused by phonemic and nonphonemic variability on early phonological encoding. More specifically, high nonphonemic variability may draw infants' attention to the phonemic invariants (i.e. similarity in phonological composition across encounters of a word), and low nonphonemic variability may draw infants attention towards the nonphonemic invariants and away from the phonemic invariants (i.e. similarity in factors such as affective prosody across encounters of a word).

One potential limitation of this study is that it fails to determine whether infants are insensitive to phonemic changes across words in this paradigm, or whether they are simply highly sensitive to affective changes across words in a way that diminishes their focus on phonemic alternations. In previous research, Werker and her colleagues (Werker, 2003; Werker & Curtin, 2005) have suggested that the allocation of resources might determine the specificity of lexical representations: Tasks that require attention to one type of mapping may detract from other types of mapping. Consistent with this reasoning, the recurrence of a particular affect may lead to greater depth of encoding along the dimension of affect and consequently, infants may have fewer resources to devote to phonological encoding. In light of this, to implement a truer test of whether infants are in fact insensitive to phonological properties of words when encountered in a single affect, it would be necessary to demonstrate that they fail to respond differentially to words and non-word minimal pairs in a within-subjects design. Furthermore, it would be important to demonstrate that infants treat words as equivalent if they match in affect, regardless of whether they match phonemically (i.e. in this scenario, infants would incorrectly equate minimal pairs if they are matched in affect). The following experiment was designed to investigate how infants treat affective equivalence in the face of subtle phonological changes and furthermore, how they treat phonological equivalence in the face of potentially more perceptible (yet lexically irrelevant) affective changes.

Experiment 3

The purpose of the following experiment was to independently manipulate phonological cues and affective cues to determine which of these sources of information infants prioritize in early word segmentation. In this experiment, all infants were familiarized with a word and a non-word (minimal pair of the target word). Therefore, lexical status was manipulated as a within-subjects factor. However, the assignment of affect to familiarization tokens (words or non-words) was manipulated between subjects: For half of the infants, the word presented during familiarization was matched in affect to the recognition passages and the non-word presented during familiarization was mismatched in affect to the recognition passages. For the other half of infants, the word presented during familiarization was mismatched in affect to the recognition passages, and the non-word was matched in affect to familiarization passages. From this design, it is possible to determine how infants treat phonological equivalence and affective equivalence and more importantly, whether they tolerate affective mismatches when words are phonologically matched and similarly, whether infants detect phonological matches that are affectively mismatched.

Participants

Thirty-two full-term, English-exposed 7.5 month olds participated in the study (13 males and 19 females), recruited from Massachusetts birth records. The mean age of participants was 241 days (range = 219 days to 236 days). Three additional infants were tested and data were discarded because of inattention.

Stimuli

Stimuli consisted of four monosyllabic non-words (“dike”, “gat”, “bree” and “gare”) and four monosyllabic word counterparts (“bike”, “hat”, “tree”, “pear”). The stimuli were identical to those used in Experiment 2 and in previous studies (Singh et al., 2004) in speaker and recording environments. Stimuli analyses are displayed in Tables 1 and 3. The passages were the same as those used in Experiments 1 and 2.

Apparatus and Procedure

Familiarization words for each infant consisted of a word and a non-word (i.e. “dike” and “hat”), where one was matched in affect and one was mismatched. The within-subjects manipulation was lexical status of the word, with all infants receiving one word and non-word. The assignment of affect to each familiarization word was counterbalanced across subjects. Half of the infants heard a non-word in an affect that matched the recognition passages and a word that mismatched the affect of recognition passage – in this condition, infants were presented with an affective match yet phonemic mismatch and a phonemic match yet affective mismatch. The other half of the infants heard a word that was matched in affect (an affective and phonemic match) and a non-word that was mis-matched in affect (an affective and phonemic mismatch). While we predict based on Experiment 2 that infants will recognize words matched in affect and fail to recognize non-word mismatched in affect, it remains unclear how infants will prioritize affective matches compared with phonemic matches when these two factors are independently controlled and manipulated as in the present experiment. In other words, in such a design, does affective equivalence or phonemic equivalence predominate in word recognition?

During the test phase, infants heard four passages. Half of the infants heard all passages in happy affect and the other half heard all passages in neutral affect. Therefore, there was one other between-subjects condition, passage affect during recognition (happy or neutral), and one within-subject condition, word affect during familiarization. The procedure was otherwise identical to that of Experiments 1 and 2.

Results and Discussion

As before, difference scores were used as an index of word recognition. In addition, within-group analyses were performed individually for infants who heard happy and neutral recognition passages, due to observed directional effects (inhibition of attention to neutral familiarization words versus increased attention to happy familiarization words).

For each recognition passage type (happy or neutral), infants were familiarized with one word and one similar sounding non-word and then tested on recognition of both words in the context of passages. In the first condition, the familiarization word was affectively matched and the familiarization non-word was affectively mismatched to the recognition passages. In the second condition, the familiarization word was affectively mismatched and the familiarization non-word was affectively matched to the recognition passages. There were therefore two factors: phonemic match and affective match. Analyses were conducted separately for happy and neutral passages comparing the effects of these two factors. For happy passages, in a 2×2 ANOVA conducted on recognition scores, infants showed a main effect of affective match,

$F(2,28) = 9.35, p < .01$. There was no main effect of phonemic match and furthermore, no significant interaction of affective and phonemic match. Planned individual comparisons were conducted within each condition. For the first condition, these analyses revealed significant recognition scores for affectively matched words (phonemic and affective match), $t(7) = 4.6, p < .01$ but non-significant recognition scores for mismatched non-words (phonemic and affective mismatch), $t(7) = -.54, NS$. Individual comparisons for the second condition revealed significant recognition scores for matched non-words (phonemic mismatch and affective match), $t(7) = 2.38, p = .05$ and non-significant recognition scores for mismatched words (phonemic match and affective mismatch) $t(7) = -.62, NS$. Thirteen of sixteen infants recognized affectively matched items (across both words and non-words) and eight of sixteen infants recognized phonemically matched items (across affectively matched and mis-matched items). These results demonstrate that infants in this condition recognized affectively matched words regardless of whether they were phonemically matched or not. Raw data are documented in Table 5 and recognition scores for each condition are graphed in Figures 6a and 6b.

A very similar pattern of results emerged in the neutral passage condition. In a similar 2×2 ANOVA, there was a significant main effect of affective match, $F(2,28) = 4.87, p < .05$. There was no main effect of phonemic match on recognition scores, nor was there any interaction of affective and phonemic matches. Planned individual comparisons for the first condition revealed significant recognition scores for affectively matched words (phonemic and affective match), $t(7) = -4.19, p < .01$ but non-significant recognition scores for mismatched non-words (phonemic and affective mismatch), $t(7) = -1.29, NS$. Pairwise comparisons for the second condition revealed significant recognition scores for matched non-words (phonemic mismatch and affective match), $t(7) = -2.83, p < .05$ and non-significant recognition scores for mismatched words (phonemic match and affective mismatch) $t(7) = .10, NS$. Twelve of sixteen infants recognized affectively matched items (across both words and non-words) and six of sixteen infants recognized phonemically matched items (across both affectively matched and mis-matched items). These results demonstrated that, as in the case of happy passages, infants who encountered neutral test passages recognized affectively matched familiarization items both when they were exact phonemic matches and when they were close phonemic matches. Conversely, infants rejected affective mismatched items both when they were phonemically matched and mismatched.

Overall, these findings demonstrate that infants showed significant recognition scores for affectively matched items, regardless of whether those items were phonemically matched or mismatched. Findings suggest that infants did not reject similar sounding nonwords that were matched in affect. Furthermore, they did not recognize words that were mismatched in affect. In sum, these findings demonstrate that infants prioritized affective equivalence over phonemic equivalence and did not respond differentially to phonemic equivalents versus phonemic near-matches. They did however respond differentially to affective matches versus affective mismatches.

This study demonstrates two negative consequences of infants' sensitivity to affect when items are presented with low variability. First, infants falsely recognized non-words that were affectively matched as well as words that were affectively matched. Second, infants failed to recognize instances of words that were affectively mismatched even though they were phonemically matched. These findings corroborate the conclusions of Experiment 2 by demonstrating that low affective variation may degrade word recognition by leading infants to categorize words based on affective similarity rather than on phonemic identity.

A final question raised by this study is whether high affective variation mitigates this degradation. In other words, when words are familiarized with high affective variation, do infants treat words and non-words more selectively, correctly recognizing the former and

rejecting the latter? Experiment 4 is designed to determine whether differential responsiveness to words and non-words can be observed when there is high affective variation in the familiarization set.

Experiment 4

The goal of this experiment was to examine a hypothesis invited by the results of Experiment 3. If low variability leads to false recognition of non-words as well as false rejection of words that differ in affect across encounters, does high variability improve infants' performance in either or both respects? More specifically, does affective variability assist infants in recognizing words that differ in surface form and furthermore, does it assist infants in correctly identifying and rejecting similar-sounding non-words? In this experiment, infants heard one word with high affective variability (as in Experiment 1) and one similar sounding non-word with high affective variability. Infants then heard both words embedded in either happy or neutral recognition passages. The goal of this experiment was to seek evidence of differential responses to both words and non-words. Experiment 3 revealed similar responses to words and non-words: Recognition of both words and non-words when they matched in affect with recognition passages as well as rejection of both words and non-words when they contrasted in affect with recognition passage. In light of this, evidence of differential treatment of words and non-words in the current experiment would presumably be attributable to high variability in the familiarization set.

Participants

Thirty-two full-term, English-exposed 7.5 month olds participated in the study (21 males and 14 females), recruited from Massachusetts birth records. The mean age of participants was 227 days (range = 211 days to 228 days). Two additional infants were tested and data were discarded because of inattention.

Stimuli

Stimuli consisted of four monosyllabic non-words (“dike”, “gat”, “bree” and “gare”) and four monosyllabic word counterparts (“bike”, “hat”, “tree”, “pear”) used in Experiment 3. Each infant heard one word and one non-word. The passages were the same as those used in Experiments 1, 2 and 3. Stimulus analyses are shown in Table 3 (variable non-words).

Apparatus and Procedure

Familiarization words for each infant were a word and a non-word (i.e. “dike” and “hat”) as in Experiment 3. However, both items were introduced with high affective variability. There was one important methodological difference between this study and Experiment 1 where words were introduced with high variability. In Experiment 1, infants were familiarized with a word in five affective styles (happiness, neutrality, fear, anger and sadness) and were then tested on recognition of these words in recognition passages spoken in either happy or neutral affect. It is possible that this design is problematic on the grounds that a subset of each familiarization set matches the affect of the familiarization passages. This limits the interpretation that variability has assisted infants in recognizing words in contrastive affect during familiarization and recognition as some of the stimuli are affective matched. Therefore, in the current experiment, the affect of the recognition passages was excluded when forming the familiarization set and reserved exclusively for testing. For example, an infant who heard recognition passages in neutral affect would have heard familiarization items in a neutral, sad, angry or fearful voice, but not in a happy voice. If infants were to recognize such a word, this would provide stronger evidence for their abilities to recognize words that differ in affect on account of high surface variation. Infants therefore heard one word in variable affect and one

non-word in variable affect, with each variable set excluding the affect of the recognition passages (either happy or neutral) for a particular infant.

During the test phase, infants heard four passages. Half of the infants heard all passages in happy affect and the other half heard all passages in neutral affect. Therefore, there was one other between-subjects condition, passage affect during recognition (happy or neutral), and one within-subject condition, lexical status (word or non-word) during familiarization. The procedure was otherwise identical to that of Experiments 1, 2 and 3.

Results and Discussion

In this experiment, infants were familiarized with one word in variable affect and one similar sounding non-word in variable affect. Infants were then tested on recognition of these items in happy and neutral passages. As in previous experiments, difference scores were used as a measure of word recognition with raw data presented in Table 5.

The experimental design comprised two factors: lexical status (word or non-word) and passage affect (happy or neutral recognition passages). A 2×2 ANOVA designed to probe the effects of these factors on recognition scores revealed a main effect of lexical status, $F(1,30) = 5.82$, $p < .05$, whereby recognition scores for words were significantly higher than for non-words. There was no main effect of passage affect (happy versus neutral), nor was there any interaction of lexical status and passage affect.

Due to directional effects observed in previous experiments for happy and neutral items, results were broken down by passage affect as in the preceding experiments. These data are graphed in Figures 8 and 9. Planned individual comparisons revealed significant recognition scores for words presented in variable affect, $t(15) = 2.73$, $p < .05$, with twelve of sixteen infants showing this effect. In this condition, there were non significant recognition scores for non-words presented in variable affect $t(15) = .82$, NS, with thirteen of sixteen infants showing no recognition of non-words. Similar results were found in the neutral passage condition, where infants showed significant recognition scores for words presented in variable affect, $t(15) = 3.8$, $p < .01$, with thirteen of sixteen infants showing this effect and non-significant recognition scores for non-words presented in variable affect, $t(15) = .55$, NS, with ten of sixteen infants showing this effect. Results by condition are depicted in Figures 8 and 9.

In summary, these findings suggest that variable familiarization enables infants to recognize words amidst differences in affect. Furthermore, it enables infants to differentiate words and similar-sounding non-words, correctly recognizing the former and correctly rejecting the latter. These findings stand in striking contrast to those of Experiment 3 where infants appeared to recognize words (and similar sounding non-words) on the basis of affective similarity. They also appeared to reject both types of stimuli on the basis of affective dissimilarity in Experiment 3.

A between-subjects comparison of Experiments 3 and 4 was conducted to determine whether infants' abilities to recognize affectively mismatched words was related to the presence or absence of variability in the familiarization set. In both experiments, the familiarization set did not match the recognition set in affective form; in Experiment 3, the familiarization set constituted a variable set of tokens and in Experiment 4, the familiarization set constituted a uniform but contrastive set of tokens compared with the recognition stimuli. Therefore, in both experiments, infants were expected to recognize a word that differed in form during familiarization. Results of a oneway ANOVA revealed a main effect of variability on recognition scores of affectively mismatched words, $F(1, 22) = 4.4$, $p < .05$ for happy passages and a marginally significant effect of variability for neutral passages, $F(1, 22) = 2.78$, $p = .10$. This suggests evidence that high variability during familiarization allows infants to recognize

words in spite of dissimilarities in affect across familiarization and recognition, with infants showing higher recognition scores when words were familiarized with high variability than when they were familiarized with low variability.

A useful corollary might be to investigate whether variability exerts an effect on infants' rejection of similar sounding non-words. Therefore, a oneway ANOVA was conducted to determine whether recognition scores for non-words varied depending on whether non-words were introduced with high or low variability. In this analysis, recognition scores for non-words presented with low variability (both affectively matched and mismatched) were compared with recognition scores for non-words presented with high variability. This analysis did not yield a significant main effect of variability for happy or for neutral passages. This finding is perhaps not surprising given that infants presumably rejected non-words in Experiment 4 based on the hypothesis that these non-words were introduced with high variability. However, infants also rejected non-words in Experiment 3 when those non-words contrasted in affect across familiarization and recognition. This finding is hypothesized to be due to an inability to recognize affectively dissimilar items in this experiment, whether those items were words or non-words. Therefore, the non-significant effect may be due to similar treatment of non-words in one condition of Experiment 3 (mismatched non-words) and Experiment 4 although the basis for this treatment is likely to be different. This differential treatment is evidenced by the fact that infants in Experiment 4 recognized affectively dissimilar words and rejected affectively dissimilar non-words, whereas infants in Experiment 3 rejected affectively dissimilar words and non-words alike. Therefore, the basis on which non-words were rejected is hypothesized to be affective dissimilarity in Experiment 3 and phonemic dissimilarity in Experiment 4.

General Discussion

The present set of experiments investigated the extent to which infants' memories for words are influenced by the variability with which those words are initially encountered. Across four experiments, infants' sensitivity to relevant (phonological) and irrelevant (affective) detail were independently manipulated and controlled in order to assess the factors that most powerfully determine success and failure at the earliest stage of word recognition. A summary of each experimental design and findings is provided in Table 4. In Experiment 1, seven- to eight-month-old infants were familiarized with a pair of target words, one of which was presented in varying forms and the other of which was presented in a uniform style. Infants were able to recognize the word introduced in varying affective forms when it was later embedded in fluent speech. Moreover, infants were able to recognize the other word which was presented in uniform affect, even though the affect of that word was mismatched across the familiarization and recognition phases of the experiment. The ability to recognize words that are mis-matched in surface form using the current methodology was previously observed only in infants at the age of 10.5 months (Houston & Jusczyk, 2000; Singh, et al., 2004). Therefore, infants' abilities to robustly recognize words in fluent speech appears to be mediated by the quantity of experience with words in their language, resulting in mature word recognition presumably as a function of simply having accumulated more instances of words by 10.5 months, but also by the quality of experience with instances of words in their language, as a function of accumulating more varied instances of words which can exert beneficial effects as young as 7.5 months.

The beneficial effects of high variability observed in Experiment 1 complement previous investigations of talker variation on infant word recognition at 7.5 months. In a similar design, Houston (1999) investigated the effects of single versus multiple talkers on the generalizability of infants' lexical representations. In a precursor to this study, Houston & Jusczyk (2000) reported that young infants were not able to equate instances of words that were spoken by a

male and by a female. However, by familiarizing infants with instances of a word spoken by multiple talkers, Houston uncovered evidence of more abstract word recognition. In this set of studies, infants were able to recognize words with greater success when they had been trained on words produced by multiple talkers of both genders than when they were trained on words produced by a single talker.

Conversely in Experiment 2, familiarization with a relatively uniform set of exemplars compromised early word recognition possibly by increasing the perceived importance and diagnosticity of surface similarity. When exemplars presented during familiarization were highly similar to one another in one aspect of surface form, such as affect, infants appeared to place a high premium on surface similarity, eclipsing their attention to subtle phonological changes. This was exemplified by infants' tendency to falsely recognize affectively matched variants that differed in the onset consonant. When placed in the context of previous investigations of false recognition where infants did not falsely recognize minimal pairs as instances of the target word (Jusczyk & Aslin, 1995; Tincoff & Jusczyk, 1996), it appears that may infants require a certain degree of surface variability to detect the kinds of subtle phonological distinctions that distinguish minimal pairs, such as "bike" and "dike". The stimuli employed by Jusczyk & Aslin (1995) were produced in typical infant-directed speech of which high affective variation is a signature trait.

The results of Experiment 3 provide more compelling evidence of infants' heavy reliance on surface form when equating encounters of words. In this experiment, infants' abilities to recognize words were determined by whether they matched in affect, regardless of whether they matched phonemically. Infants treated non-words and words alike, recognizing both stimulus items when they matched in affect and failing to recognize both stimulus items when they mismatched in affect. This provides strong evidence that infants prioritize surface form in early word recognition at the cost of relevant phonological detail². By contrast, the findings of Experiment 4 demonstrate a greater sensitivity to phonological detail when words are introduced with affective variation: Infants recognize words familiarized in variable affect and fail to recognize non-words familiarized in variable affect. Therefore, unlike in Experiment 3, infants treat words and non-words differentially in the context of high affective variation. Combining the results of Experiment 1 (words familiarized with high and low affective variation) and Experiment 4 (words and non-words familiarized with high affective variation), it would appear that affective variability has dual benefits for early word recognition. First, it assists infants in identifying the phonological invariants for the variable word as well as other words presented in the same context, allowing infants to disregard irrelevant surface details. Second, it assists infants in discriminating phonological matches and near-phonological matches leading to a greater degree of phonological specificity in early speech segmentation. Therefore, it appears that high affective variation has had the effect of enhancing infants' perception and retention of invariant phonological detail, possibly inviting another of several attested advantages conferred by infant-directed speech. Accordingly, it is possible that infants' perceptual acuity for phonological detail is optimized for the intrinsically varying cadence of infant-directed speech.

²There are at least two possible ways to account for this finding. First, it is possible that infants, by default, prioritize dimensions such as vocal affect when equating words by virtue of its social/emotional appeal. By extension, unless they receive counterevidence that vocal affect varies orthogonally to lexical identity, they may preferentially match words based on affect, thereby automatically weighting affective quality more heavily than the identity of phonetic segments. Another possibility is that infants truly become insensitive to segmental constancy when matching words based on suprasegmental constancy because their attention to more transient acoustic events is compromised when more enduring events, such as affect, are emphasized by the task. The current findings do not distinguish between these possible mechanisms and certainly, both are plausible accounts. Cross-linguistic evidence probing these effects in tonal languages when words differ phonemically in lexical tone versus in consonant identity may clarify this issue.

Positive effects of between-talker variability have been observed on word segmentation and on mapping novel words to meaning (Houston, 2000; Hollich, Jusczyk & Brent, 2002). Using a different source of variability (different talkers across both genders), both Houston and Hollich et al., have revealed difficulty on the part of preverbal infants (7.5 months) and older children (24 months) in generalizing across talkers when presented with unknown words by a single talker. However, both investigators have independently demonstrated that when infants and children are introduced to a word by several talkers, their ability to generalize across talkers is greatly enhanced allowing for effective word segmentation at 7.5 months as well as effective word to object mapping at 24 months. These results are compatible with the present findings in that collectively, they implicate high variability as a potential cue to discovering the phonological invariants of a word amidst conditions of high uncertainty (i.e. when infants do not know the meaning of the word).

In summary, varying degrees of abstraction and specificity can emerge as a consequence of altering the distributional profile of words presented to infants and children. When introduced by words in varying forms, learners appear to expand the range of exemplars that they will admit to that lexical category and appear to correctly reject exemplars that do not fall into that lexical category either because they are affectively mismatched (as in Experiment 1) or because they are phonemically mismatched (as in Experiment 4). When they are introduced to words with minimal variation, they appear to only admit highly similar exemplars to that category, forming a comparatively narrow lexical category. At the broadest level, these findings argue for dynamic memory representations in early learners that can be automatically reconstituted depending on the nature of the input. Specifically, in certain test situations, infants demonstrate evidence of highly specific lexical representations, whereas in other test situations, infants appear to define words by their abstract properties. This versatility may allow them to flexibly represent properties words and their defining characteristics online, in a way that is attuned either to exemplar- or to category-specific detail, depending on the level of processing cued by the task.

This account suggests that infants are capable of retaining in memory fine detail about words they hear and that they match encounters of a word in a task-dependent manner. If infants are capable of encoding both phonological and perceptual details, why does perceptual similarity reduce sensitivity to phonological details as observed in Experiments 2 and 3? Why should one dimension of sound interfere with the other if infants have the capacity to encode variability along both dimensions? This issue is usefully informed by results of adult speech processing tasks. Studies investigating adult lexical representations have repeatedly demonstrated that phonological and perceptual details are encoded in a manner that is highly interdependent (Mullenix & Pisoni, 1990; Palmeri, Goldinger & Pisoni, 1993). It is not the case that perceptual and phonological details are carefully partitioned into two separate pathways by the nervous system, rather, they are mutually dependent sources of information (Goldinger, 1998). When asked to process phonological detail or perceptual detail in speech, changes across one dimension interfere with processing of the other dimension. In other words, when asked to classify voices, listeners' performance deteriorates if the stimuli vary phonologically and vice versa (Mullenix & Pisoni, 1990). The issue of whether perceptual and phonological variation are mutually interfering is relevant to the present investigations because in previous studies (Houston & Jusczyk, 2000; Singh et al., 2004), infants showed matching effects in which the presence of perceptual variation interfered with their ability to classify tokens based on phonological characteristics. In other words, they were unable to equate two phonologically equivalent instances of "bike" because they varied perceptually. However, if this interference were symmetrical, we would expect changes in phonological structure to interfere with classification along perceptual dimensions. On the contrary, in Experiment 2, when infants were presented with phonological variation (e.g. "bike" and "gike"), they did classify tokens along perceptual dimensions, matching stimuli based on affect, suggesting that infants formed

perceptually equivalent classes in spite of phonological differences. Perhaps surprisingly, adults show a very similar asymmetry, even though they have considerable experience with words and the wide variety of forms they can assume. When asked to classify tokens based on perceptual characteristics, such as pitch, speech rate and other talker-specific traits, listeners find it relatively easy to disregard phonological changes, such as phonemic alternation. In contrast, when asked to classify instances based on phonological characteristics, listeners' performance is severely disrupted by changes in perceptual detail (Mullenix & Pisoni, 1990). These findings implicate effects of perceptual salience in both infancy and adults, whereby voice details are prominent in the speech percept and therefore may take precedence over processing of relatively subtle phonological details. These effects are observable in adults, even when they have explicit awareness of the changes they have been instructed to monitor and metalinguistic knowledge of the rules linking word tokens and lexical types. However, infants are equipped neither with direct instruction about where to focus their attention nor with the requisite knowledge of how to derive types from tokens, and are therefore possibly even more inclined to attend to certain dimensions of sound at the expense of attention to phonological detail.

Clearly, an inattention to relevant phonological detail, even one that is context-dependent, would come at a severe cost to the process of language acquisition. Therefore, prior to establishing a mature lexicon, infants must be sensitive to phonemic detail in a way that permits an accurate and precise mapping of forms (words) to meaning. Single-feature differences between words can distinguish meaning and therefore, must be incorporated into lexical representations. Studies investigating infants' sensitivities to differences in the onset consonant of a word in a word learning context have revealed strong effects of word familiarity on phonological sensitivity (Fennell & Werker, 2003; 2004; Stager & Werker, 1997; Swingley & Aslin, 2000), whereby infants successfully heed subtle cues to phonemic identity when they are familiar with at least one of the words in a minimal pair. When both words are novel, as demonstrated by Stager and Werker (1997), infants are surprisingly insensitive to changes in the onset consonant of a word. Collectively, these findings demonstrated that infants at the same age group display quite different sensitivities depending on the task demands. In a word learning situation, when infants are charged with the task of mapping novel forms to novel objects, this seems to reduce their sensitivity to phonological detail, even that to which they are assuredly sensitive in discrimination tasks. Therefore, it is possible that the task of mapping novel words to novel objects taxes resources considerably and to an extent that compromises the phonological precision with which infants are able to represent these forms in memory.

In the present set of studies, infants demonstrate differential sensitivities to similar phonological distinctions depending on the task conditions, recognizing such distinctions in conditions of high variability and appearing not to do so under conditions of low variability. Such a performance differential is usefully captured by Werker and Curtin's (2005) developmental framework of infant speech processing (PRIMIR). Central to this framework is the notion that infants will access and prioritize different types of information depending on the task at hand, revealing different sensitivities in different experimental contexts. The present findings are consistent with this framework in that infants' performance seems to depend on the particular set of cues primed by the task. When words vary in affect, this may serve to downplay the reliability with which affect can be associated with a word, thereby direct infants' attention to properties that are reliably associated with a word. When words are constant in affect, infants' tendencies to be engaged by vocal affect combined with its recurrence in the training set may eclipse their sensitivity to more subtle phonological detail. Therefore, the present set of findings

is best encapsulated by a theoretical framework that allows for a keen responsivity on the part of learners to changes in task demands, salience levels, and contexts. Additionally, these

findings, like PRIMIR, reveal the inevitable constraints of a finite set of resources. As a result, when attentional resources are consumed by a particular property (i.e. affect), they may be less available to process other properties of speech (e.g. phoneme identity). While this may initially appear to point to an unstable learning apparatus, it is in fact quite adaptable for learners who cannot arrive preconceived for the particulars of any given language. Rather, these particulars must be detected and extracted from the linguistic environment and analyzed in the service of underlying structure and organization. This initial state of uncertainty demands perceptual flexibility, the capacity for re-organization and a sensitivity towards the relative weighting of cues in the input in order to a linguistic system to mature and crystallize.

Investigating the effects of variability has consequences not only for the process of language learning but also for modeling the structure of the learning apparatus, specifically, the emergent mental lexicon. Traditional views of the mental lexicon assumed that we store words and their constituents in canonical form, discarding irrelevant surface detail (Blumstein & Stevens, 1980; Fant, 1960; Tenpenny, 1995). The present study contributes to a burgeoning store of evidence that words in fact are not stored in abstract terms, but rather, that both infants and adults incorporate surface detail and phonemic properties in memory (Schacter & Church, 1992; Goldinger, 1996; 1998; Houston, 1999; Ryalls & Pisoni, 1997). It is clear that infants do not arrive at the task of word recognition equipped with a prototypical, idealized set of lexical representations that would afford successful recognition across dissimilar instances of words. Instead, they appear to store acoustic and phonetic detail with remarkable precision and at 7.5 months, appear undecided as to how different cues ought to be weighed. These findings most readily lend themselves to an exemplar model of lexical categorization that presupposes storage of instances in a way that mirrors their veridical form. However, from observing the effects of high and low variability on infant word recognition, it appears that infants exploit the degree of dispersion between highly variable exemplars to venture predictions about acoustic-phonetic cues to lexical relevance. They then appear to apply these predictions in a wholesale fashion to other lexical types presented in the same test session.

Traditionally, variability effects in speech perception have been construed as disruptive influences on the speed and accuracy with which words are processed (Mullenix, Pisoni & Martin, 1989; Palmeri, Goldinger & Pisoni, 1990; Schacter & Church, 1992). However, it appears that variability effects can lead to improvements in performance in particular circumstances, in particular, in conditions of high uncertainty. Certainly, the ability to distill relevant cues from the abundance of cues transmitted through speech is crucial to word recognition. However, it remains unclear from this set of studies how variability interacts with lexical development when words carry meaning for infants. It is possible that infants' lexical categories are strengthened by the acquisition of meaning, at which juncture surface variation may not assist word recognition to this degree. In fact, high surface variation may actually delay or disrupt infant word recognition for items that are meaningful to infants. Therefore, it is important to identify whether the effects of variability observed here endure through lexical development or whether they serve as temporary workarounds for sorting items that carry no conceptual definition. This issue is presently being investigated in follow-up studies. For example, while infants at 7.5 months do not know the meaning of words such as 'bike' and 'dog', they do have referential knowledge of their own names (Mandel, Jusczyk & Pisoni, 1996) and words by which they know their mother and father (Tincoff & Jusczyk, 1999). Therefore, these are possible stimuli with which to disentangle the interaction of variability and lexical status in infants of a comparable age. Finally, it is important to determine whether the beneficial effects of variability observed in the present studies have enduring consequences for word recognition or whether they simply represent short-term priming effects that are limited to the task at hand. It has been found that infants retain words they were familiarized with for as long as two weeks using this procedure (Jusczyk & Hohne, 1997). Given this, it would be interesting to compare infants' long term memories of words to which they were

introduced with variation to those that they encountered without variation. This too is currently being investigated as a follow-up study to the present set of findings.

To summarize, a formidable challenge to early word recognition is the issue of how to transcend the infinite amount of irrelevant variability in speech and arrive at a stable, accurate set of lexical representations. Infants appear to embark on this process several months before they attach meaning to words. The present set of studies suggests that they may induce the properties used to define words by comparing the frequency with which different cues covary across encounters with a word, a mechanism perhaps common to the origination of non-linguistic categories in infants and adults. When cues covary perfectly across encounters, infants may assume that those cues are relevant to lexical identity. When certain cues fail to covary across encounters, infants may ascribe secondary importance to those cues. In this way, the development of an early lexicon may actually be consolidated rather than thwarted by the natural variability in everyday speech, and strengthened in particular by the enhanced variability of infant-directed speech. While it remains unclear whether these strategies are disengaged when words begin to assume meaning, infants' attention to both phonemic and non-phonemic variability in surface form appears to constitute one possible mechanism by which they accrue knowledge about the defining characteristics of words.

Acknowledgments

A section of this work formed part of a doctoral dissertation submitted by L.S. to the Brown University Graduate School and was supported by a Dissertation Fellowship from Brown University to L.S. and a grant from the National Institutes of Health (1 RO1 HD32005) to James Morgan. Other sections were supported by a grant from the National Institutes of Health to Leher Singh (5R03HD046676-02). I am very grateful to Jim Morgan, Katherine Demuth and Cathi Best for valuable feedback on this work, to Sarah Nestor, Karen Rathbun and Lori Rolfe for testing infants and to Crystal Wilson and Wendy Zosh for stimulus recording.

REFERENCES

- Anderson J, Morgan J, White K. A Statistical Basis for Speech Sound Discrimination. *Language and Speech* 2003;46(2–3):155–182. [PubMed: 14748443]
- Banase R, Scherer KR. Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology* 1996;70(3):614–636. [PubMed: 8851745]
- Best, CT. A direct realist view of cross-language speech perception.. In: Strange, W., editor. *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-language Speech Research*. York; Timonium, MD: 1995.
- Blumstein SE, Stevens KN. "Perceptual invariance and onset spectra for stop consonants in different vowel environments". *Journal of the Acoustical Society of America* 67 1980;2:648–662.
- Benedict H. Early lexical development: comprehension and production. *Journal of Child Language* 1979;6:183–200. [PubMed: 468932]
- Fant, G. *Acoustic theory of speech production*. Mouton; The Hague, Netherlands: 1960.
- Fennell CT, Werker JF. Early word learners' ability to access phonetic detail in well-known words. *Language and Speech* 2003;46:245–264. [PubMed: 14748446]
- Fernald A. Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development* 1993;64:657–674. [PubMed: 8339687]
- Gaskell MG, Marslen-Wilson WD. Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance* 1996;22:144–158. [PubMed: 8742258]
- Goldinger SD. Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1996;22:1166–1183.
- Goldinger SD. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 1998;105(2):251–279. [PubMed: 9577239]
- Gow DW. Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language* 2001;45:133–159.

- Gow DW. Does English coronal place assimilation create lexical ambiguity? *Journal of Experimental Psychology: Human Perception and Performance* 2002;28:163–179.
- Hollich, G.; Jusczyk, P.; Brent, M. How infants use the words they know to learn new words.. Paper presentation at Bstn University Conference on Language Development.; Boston, MA. 2000.
- Houston, DM. The Role of Talker Variability in Infant Word Representations. Unpublished doctoral dissertation. Johns Hopkins University; Baltimore: 1999.
- Houston DM, Jusczyk PW. The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance* 2000;26(5):1570–1582. [PubMed: 11039485]
- Jusczyk PW, Aslin RN. Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology* 1995;29:1–23. [PubMed: 7641524]
- Jusczyk PW, Hohne EA. Infants' memory for spoken words. *Science* 1997;277:1984–1986. [PubMed: 9302291]
- Kemler Nelson D, Hirsh-Pasek K, Jusczyk PW, Wright Cassidy K. How the prosodic cues in motherese might assist language learning. *Journal of Child Language* 1989;16:55–68. [PubMed: 2925815]
- Kitamura, C.; Burnham, D. The infant's response to maternal vocal affect.. In: Rovee-Collier, C.; Lipsitt, L.; Hayne, H., editors. *Advances in infancy research*. 12. Stamford, CT; Ablex: 1998. p. 221-236.
- Klatt, DH. Review of selected models of speech perception.. In: Marslen-Wilson, W., editor. *Lexical representation and process*. MIT Press; Cambridge: 1989. p. 169-226.
- Madole KL, Cohen LB. The role of object parts in infants' attention to form-function correlations. *Developmental Psychology* 1995;31:637–648.
- Madole KL, Oakes LM. Making sense of infant categorization: Stable processes and changing representations. *Developmental Review* 1999;19:263–296.
- Mandel DR, Jusczyk PW, Pisoni DB. Infants' recognition of the sound patterns of their own names. *Psychological Science* 1995;6:315–318.
- Maye J, Werker JF, Gerken LA. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 2002;82(3):B101–B111. [PubMed: 11747867]
- Mullennix JW, Pisoni DB. Stimulus variability and processing dependencies in speech perception. *Perception & Psychophysics* 1990;47:379–390. [PubMed: 2345691]
- Mullennix JW, Pisoni DB, Martin CS. Some effects of talker variability on spoken word recognition. *Journal of Acoustical Society of America* 1989;85:365–378.
- Needham, A.; Dueker, GL.; Modi, AC. Memories of objects influence infants' object segregation.. Paper presented at International Conference on Infant Studies; Toronto, ON, Canada. 2002.
- Palmeri TJ, Goldinger SD, Pisoni DB. Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1993;19:309–328.
- Pisoni, DB. Some thoughts on “normalization” in speech perception.. In: Johnson, K.; Mullennix, JW., editors. *Talker Variability in Speech Processing*. Academic Press; San Diego, CA: 1997.
- Ryalls BO, Pisoni DB. The effect of talker variability on word recognition in preschool children. *Developmental Psychology* 1997;33:441–451. [PubMed: 9149923]
- Schacter DL, Church BA. Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 1992;18:915–930.
- Scherer K. Vocal affect expression: A review and a model for future research. *Psychological Bulletin* 1986;9:143–165. [PubMed: 3515381]
- Singh L, Morgan J, Best C. Infants' Listening Preferences: Baby Talk or Happy Talk? *Infancy* 2002;3(3):365–394.
- Singh L, White K, Morgan. Building a lexicon in the face of phonetic variability: Effects of pitch and amplitude variation on early word recognition. *Language Learning and Development*. (in press)
- Singh L, Morgan J, White K. Preference and processing: The role of speech affect in early spoken word recognition. *Journal of Memory and Language* 2004;51(2):173–189.
- Stager CL, Werker JF. Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature* 1997;388:381–382. [PubMed: 9237755]

- Swingle D, Aslin RN. Spoken word recognition and lexical representation in very young children. *Cognition* 2000;76:147–166. [PubMed: 10856741]
- Tenpenny PL. Abstractionist versus episodic theories of repetition priming and word identification. *Psychonomic Bulletin & Review* 1995;2:339–363.
- Tincoff, R.; Jusczyk, PW. Are word-final sounds perceptually salient for infants?. Paper presented at the Fifth Conference on Laboratory Phonology; Evanston, IL. 1996.
- Trainor LJ, Austin CM, Desjardins RN. Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science* 2000;11:188–195. [PubMed: 11273402]
- van der Weijer, J. Language input for word discovery. MPI Series in Psycholinguistics (No. 9); Nijmegen, The Netherlands. 1998.
- Werker, JF. Phonetic encoding in early word learning.. Invited presentation at Society for Research in Child Development; Tampa, FL. 2003.
- Werker JF, Tees R. Cross-language speech perception -- evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 1984;7:49–63.
- Werker JF, Tees RC. Influences on infant speech processing: Toward a New Synthesis. *Annual Review of Psychology* 1999;50:509–535.
- Werker JF, Curtin S. PRIMIR: a developmental framework of infant speech processing. *Language Learning and Development* 2005;1:197–234.
- Williams CE, Stevens KN. Emotions and speech: Some acoustical correlates. *Journal of the Acoustical Society of America* 1972;52:233–248.

Figure 1a.
Mean fundamental frequency for variable, happy and neutral words (measured in Hertz)

Figure 1b.
Fundamental frequency range for variable, happy and neutral words (measured in Hertz)

Figure 2. Looking times (means and standard errors) for happy passages containing words familiarized in variable and constant (neutral) affect

■ Variable Familiarization Item □ Happy Familiarization Item

Figure 3. Looking times (means and standard errors) for neutral passages containing words familiarized in variable and constant (happy) affect

Figure 4. Looking times (means and standard errors) for happy passages containing words familiarized in happy and neutral affect

■ Happy Familiarization Item ■ Neutral Familiarization Item

Figure 5. Looking times (means and standard errors) for neutral passages containing words familiarized in happy and neutral affect

■ Happy (matched) word ■ Neutral (mismatched) non-word

Figure 6a.

Looking times (means and standard errors) for happy passages containing happy words and neutral non-words

Figure 6b.
Looking times (means and standard errors) for happy passages containing happy non-words and neutral words

Figure 7a.
Looking times (means and standard errors) for neutral passages containing neutral words and happy non-words

Figure 7b.

Looking times (means and standard errors) for neutral passages containing neutral non-words and happy words.

Figure 8. Looking times (means and standard errors) for happy passages containing variable words and variable non-words.

Figure 9. Looking times (means and standard errors) for neutral passages containing variable words and variable non-words.

Table 1

Acoustic Analyses of Words: Means and (Standard Deviations)

	Mean F ₀	Min F ₀	Max F ₀	F ₀ range
Happy	320.45 (48.07)	231.89 (43.66)	404.77 (39.21)	172.88 (52.14)
Neutral	158.44 (6.39)	141.64 (9.24)	184.20 (14.17)	42.56 (18.05)
Variable	256.47 (94.47)	153.86 (83.35)	473.45 (105.44)	319.57 (78.98)

Table 2

Acoustic Analyses of Words in Sentences: Means and (Standard Deviations)

	Mean F_0	Min F_0	Max F_0	F_0 range
Happy	223.64 (21.52)	145.43 (43.15)	334.57 (18.54)	189.14 (45.2)
Neutral	164.75 (27.48)	104.55 (13.63)	185.75 (5.76)	81.2 (17.64)

Table 3

Acoustic Analyses of Non-words: Means and (Standard Deviations)

	Mean F ₀	Min F ₀	Max F ₀	F ₀ range
Happy	301.13 (72.01)	175.31 (38.35)	420.06 (122.79)	244.75 (118.41)
Neutral	190.31 (18.16)	147.41 (6.38)	223.16 (19.62)	75.75 (15.14)
Variable	287.45 (128.87)	139.8 (112.84)	412.68 (146.73)	272.88 (149.46)

Table 4

Stimulus Manipulations in Each Experiment

Experiment Number	Familiarization Stimulus 1	Familiarization Stimulus 2	Recognition of Stimulus 1	Recognition of Stimulus 2
Experiment 1	Variable Word	Mismatched Word	Yes	Yes
Experiment 2	Matched Non-Word	Mismatched Non-Word	Yes	No
Experiment 3 Condition 1	Matched Non-Word	Mismatched Word	Yes	No
Experiment 3 Condition 2	Mismatched Non-Word	Matched Word	No	Yes
Experiment 4	Variable Non-Word	Variable Word	No	Yes

Table 5

Mean Looking Times (S.E.) for Each Experiment

Experiment Number	Type of Word in Passage	Happy Test Passages	Neutral Test Passages
Experiment 1	Variable Word	9866.14(911.2)	8154.54(1098.3)
	Mismatched Word	6464.26(610.25)	9598.72(1373.54)
	Unfamiliar Passages	8482.54(662.24)	6257.02(761.25)
Experiment 2	Matched Non-Word	8068.25(647.71)	5858.64(362.94)
	Mismatched Non-Word	6528.28(372.94)	6662.56(642.84)
	Unfamiliar Passages	5496.65(329.15)	7052.58(475.95)
Experiment 3 Condition 1	Matched Non-Word	11051.44(889.64)	7990.71(591.11)
	Mismatched Word	8019.34(1430.75)	9952.48(721.58)
	Unfamiliar Passages	8578.54(814.25)	9791.25(871.24)
Experiment 3 Condition 2	Mismatched Non-Word	7830.54(904.27)	7878.94(303.54)
	Matched Word	11658.54(712.97)	6246.54(410.17)
	Unfamiliar Passages	8472.9(755.44)	8389.84(414.67)
Experiment 4	Variable Non-Word	7802.18(651.88)	7524.46(597.52)
	Variable Word	9038.21(732.84)	8968.12(661.87)
	Unfamiliar Passages	7131.46(523.45)	7076.24(495.22)