



Published in final edited form as:

Cogn Psychol. 2010 June ; 60(4): 241–266. doi:10.1016/j.cogpsych.2010.01.002.

Is Early Word-form Processing Stress-full? How Natural Variability Supports Recognition

Heather Bortfeld and

Department of Psychology, University of Connecticut

James L. Morgan

Department of Cognitive and Linguistic Sciences, Brown University

Abstract

In a series of studies, we examined how mothers naturally stress words across multiple mentions in speech to their infants and how this marking influences infants' recognition of words in fluent speech. We first collected samples of mothers' infant-directed speech using a technique that induced multiple repetitions of target words. Acoustic analyses revealed that mothers systematically alternated between emphatic and nonemphatic stress when talking to their infants. Using the headturn preference procedure, we then tested 7.5-month-old infants on their ability to detect familiarized bisyllabic words in fluent speech. Stress of target words (emphatic and nonemphatic) was systematically varied across familiarization and recognition phases of four experiments. Results indicated that, although infants generally prefer *listening* to words produced with emphatic stress, *recognition* was enhanced when the degree of emphatic stress at familiarization matched the degree of emphatic stress at recognition.

Early Word Recognition may be Stress-full

Learning to recognize spoken words is a formidable task. Infants are bombarded with words that vary phonetically and acoustically across a broad range of phonological, syntactic, and discourse contexts. Other features of language vary with changes in speaker identity and affect. Adding to the challenge, infants acquiring their first language(s) must learn to group independent instances of a word (i.e., word tokens) into single categories (i.e., word types), and to do this in "real time," at the moment the utterance occurs in a given language context. Further complicating matters, words are not typically separated by pauses in fluent speech, and the cues that may serve to signal word boundaries between words vary from language to language.

Despite these challenges, infants normally begin to recognize words at around 6 months of age. This is demonstrated by their early ability to recognize highly familiar items even when embedded in a stream of speech (Bortfeld, Morgan, Golinkoff, & Rathbun, 2005; Mandel-Emer, 1997; Singh, Nestor, & Bortfeld, 2008). Over the coming weeks, they become increasingly attuned to patterns of their native language, so that by about 7.5 months of age they can recognize newly familiarized monosyllabic words within fluent speech (Jusczyk &

Corresponding Author: Heather Bortfeld, Tel: 860-486-0919, Fax: 860-486-2760 Department of Psychology, University of Connecticut, Storrs, CT 06269-1020, heather.bortfeld@uconn.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.

Aslin, 1995) and even certain types of bisyllabic words (Jusczyk, Houston, & Newsome, 1999).

The nature of infants' early lexical representations and, hence, the properties of the input speech that facilitate or hinder word segmentation and recognition are not yet well understood. Early findings suggested that infants' representations might be detailed and surprisingly adult-like. For instance, Jusczyk and Aslin (1995) found that 7.5-month-olds resemble adults in that they can recognize familiarized words in fluent speech and do not false alarm to items differing by only a single phonetic feature from familiarized targets. Subsequent studies have shown, however, that infants' word recognition can be disrupted by changes in dimensions that would be lexically irrelevant to adults, such as talker gender (Houston & Jusczyk, 2000), speaker affect (Singh, Morgan, & White, 2004), or pitch (Singh, White, & Morgan, 2008). A range of studies have shown that infants rely heavily on lexical stress (i.e., when the syllables of a multisyllabic word are not stressed equally) for assistance in spoken word recognition, a factor that influences adult lexical processing as well (Cutler & Butterfield, 1992; Cutler & Norris, 1988; Mattys & Samuel, 1997; Norris, McQueen, & Cutler, 1995; Slowiaczek, 1990; Small, Simon, & Goldberg, 1988). Here we explore how another form of stress, emphatic stress (i.e., acoustic emphasis used to call attention to a word in a given context) is used in speech directed towards infants and how its use affects early word recognition.

Influences on Early Word Recognition

To test infants' sensitivity to acoustic features such as lexical stress, researchers frequently rely on the headturn preference procedure (HPP). In the HPP (see Jusczyk & Aslin, 1995; Kemler Nelson, Jusczyk, Mandel, Myers, et al., 1995) infants are familiarized with particular words. They then hear sets of sentences that are concatenated into fluent strings of speech. Some of the sentence sets contain repetitions of the familiarized words and others contain unfamiliarized words. Infants' continued orientation towards a blinking light results in sentence sets being played; this orientation time is then measured. If infants display consistent differences in orientation times to sentences with previously familiarized versus unfamiliarized words, it may be inferred that infants have formed some representation of the familiarized items in memory and that they are able to recognize those representations (i.e., the sound patterns of the familiarized words) within that running speech.

Using this technique, Jusczyk, Houston, and Newsome (1999) explored infants' detection of bisyllabic words in fluent speech. Across a series of experiments testing 7.5- and 10-month-old English-exposed infants, they found that the younger infants were able to recognize bisyllabic words with strong-weak, but not weak-strong lexical stress patterns (and see Echols, Crowhurst, & Childers, 1997; Morgan, 1996; Morgan & Saffran, 1995). That is, the younger infants were able to recognize familiarized words with the more canonical form of lexical stress in their native language (in this case English, in which initial syllables are typically stressed in bisyllabic words), but not those with the less common pattern (in this case, in which second syllables are stressed). The older infants, on the other hand, could recognize familiarized words with either pattern of stress. Related studies have shown that lexical stress overrides cues to word boundaries from sequential statistics (Johnson & Jusczyk, 2001; cf. Saffran, Aslin, & Newport, 1996) and can alter interpretation of phonotactic patterns (Mattys, Jusczyk, Luce, & Morgan, 1999). The role that lexical stress plays in infant speech segmentation has been found for other languages with rhythms similar to English as well (for example, in German (Höhle & Weissenborn, 1999), and in Dutch (Houston, Jusczyk, Kuijpers, Coolen & Cutler, 2000)), though not for languages with different rhythms (for example, French (Nazzi, Iakimova, Bertoncini, Fredonie, & Alcantara, 2006)). Thus, with infants, as with both normal adult populations (Norris et al., 1995; Slowiaczek, 1990; Small et al., 1988) and adults with left-

hemisphere-damage (Baum, 2002), lexical stress can be represented in the mental lexicon and seems to play an important role in guiding lexical access, at least in some languages.

Other Influences: Emphatic Stress and Repetition

Although lexical stress appears to be influential in a subset of the world's languages, there are other forms of stress that might influence infant speech recognition more generally. In particular, emphatic stress can be used to signal the informational status of a word (Chafe, 1976), to signal information that is not shared between a speaker and a listener (Fowler, 1988; Solan, 1980), or to indicate the focus of a sentence (Rochemont & Culicover, 1990). In conversation between adults, the informational status of a word's referent in discourse influences how that word is produced (Cutler, Dahan, & van Donselaar, 1997; Shattuck-Hufnagel & Turk, 1996). Speakers emphasize certain words over others—typically those just introduced into the conversation—by using some combination of increased pitch, higher intensity, and longer duration. This “new” stress serves to foreground a word from the rest of the words in the utterance. After a single use of new stress, speakers generally switch to reduced (or “given”) stress (Fowler, 1988; Fowler & Housum, 1987; Fowler, Levy, & Brown, 1997; though see also Gravano & Hirschberg, 2006). In subsequent mentions, the word may be replaced by a pronoun or it may be elided entirely. The reduction of stress from initial to subsequent mentions is one component of a larger phenomenon in adult-directed speech (ADS) referred to by some researchers as the “given/new contract” (Chafe, 1976; Clark & Haviland, 1977; Halliday, 1967; Haviland & Clark, 1974; Prince, 1981). Accordingly, listeners come to expect that words receive full stress when they are initially introduced and will be produced with reduced stress in subsequent mentions (Bock & Mazzella, 1983; Cutler, 1990; Nootboom & Kruyt, 1987; Terken & Nootboom, 1987). But reduction of stress is not an all-or-none matter. Repeated words are less likely to be reduced if they are central to the topic of conversation (Fowler & Housum, 1987). Grammatical role also affects the likelihood of reducing stress on repeated words (Terken & Hirschberg, 1994). For example, Dahan and colleagues (Dahan, Tanenhaus, & Chambers, 2002) argued that adult listeners preferentially interpret accented nouns as referring not necessarily to a new item, but to a previously mentioned item that was not the focus of the immediately preceding sentence but that has become the focus of the current sentence. In other words, a growing body of research on this issue is offering a nuanced view of how even adults speaking to other adults use stress to mark information accessibility in discourse. Yet, despite the variability in the application of emphatic stress in ADS, the given/new account has remained a useful guideline for understanding how stress is used in discourse.

Interestingly, the exaggeration of word duration, pitch height, and pitch range characteristic of speech directed to infants facilitates their word learning by maximizing the salience and (perhaps) intelligibility of novel forms in much the same way that new, or emphatic, stress is used to draw adult listeners' attention to particular words. Infant-directed speech (IDS) has been identified as a form of “hyperspeech” that is used virtually universally (Fernald, 2000) and, consistent with this view, some research (e.g., D'Odorico & Jacob, 2006) indicates that lack of exposure to IDS is associated with delayed speech development in young children. However, despite their overall acoustic exaggeration when speaking to infants, speakers still tend to position novel words on pitch peaks and at the ends of utterances (Aslin, Woodward, LaMendola, & Bever, 1996; Fernald, 2000; Fernald & Mazzie, 1991), thereby emphasizing certain words beyond the emphasis already inherent in the IDS form. In this way, caretakers can still guide infants' attention to a particular word by acoustically highlighting it relative to the rest of an utterance and in a way that is similar to that observed in ADS.

One of the important features in ADS is that after an initial (usually emphatically stressed) mention and (perhaps) subsequent mention without emphatic stress, a word may be replaced

by a pronoun or it may be elided entirely. In IDS, however, an important aid to early word learning is that nouns are repeated in their full form, without being replaced with the corresponding pronoun (Ferguson, 1964). This is quite distinct from the reduction that typically occurs across mentions in ADS (Bolinger, 1972; Ladd, 1980, 1996; Selkirk, 1984). The following utterance from a mother to her 13-month-old infant (quote from Bernstein Ratner, 1996) demonstrates this:

M: I'll go get your **block!** This's a **block.** Say...mommy...**block!** Here. Ok. Now what?

This continues for several utterances, such that the word “block” is repeated to the infant at least ten times. Repetition of the word without its replacement with the corresponding pronoun (e.g., “it”) would be unexpected in competent adult conversation, but it seems perfectly normal for someone speaking to an infant. This kind of massed provision of multiple examples of words in a limited time is likely to facilitate infants’ formation of robust lexical representations, thereby enhancing recognition. But how do the two forms of perceptual highlighting, emphatic stress and word repetition, interact in IDS? The characterization of word stress in ADS as following a given/new contract has itself been challenged (e.g., Dahan, Tanenhaus, & Chambers, 2002). It is even less clear how such stress interacts with speech in which content words are repeated in full form not once or twice, but multiple times. In considering this question, it is useful to consider first whether the normal pattern of stress observed across first and second repetitions of words in ADS does, in fact, occur in IDS.

Emphatic Stress in Infant-Directed Speech

The use of emphatic stress in IDS has been identified as an important contributor to early language development (Bernstein & Ratner, 1996). There is also substantial evidence to suggest that the emphatic stress mothers use when talking to their infants parallels the acoustic structure of emphatic stress in speech between adults (e.g., Cooper & Aslin, 1990; Fernald, 1984; Fernald & Kuhl, 1987; Fernald & Mazzie, 1991; Fernald & Simon, 1984; Fernald, Taeschner, Dunn, Papousek, Boysson-Bardies, & Fukui, 1989; Werker & McLeod, 1989), and that this influences infants’ word segmentation abilities (Cutler & Swinney, 1987; Thiessen, Hill, & Saffran, 2005). However, there is much less research examining how emphatic stress combines with the multiple mentions of words typical of IDS. Documented high levels of repetition in this register (e.g., Bernstein Ratner, 1996) suggest that the use of emphatic stress may manifest differently, thereby allowed caretakers to assist in the development of word recognition by making repeated words perceptually prominent over time rather than following the pattern of reduction characteristic of ADS (e.g., emphatic stress-to-reduced stress-to-pronominalization).

There is at least some evidence that adults speak to infants using emphatic stress in much the same way they do in speech to other adults. Fisher and Tokura (1995) used a puppet-show task to elicit a designated set of target words while mothers produced spontaneous speech to their infants. The speech data were then examined to determine how emphatic stress interacted with the prosodic modifications typical of infant-directed speech. Fisher and Tokura found that mothers attenuated stress, as measured by pitch and duration changes, on the second mention of a previously mentioned word to the same degree that they did when addressing another adult, leading them to conclude that something akin to the given/new contract is observed in IDS as well. Crucially, however, Fisher and Tokura’s analyses did not look beyond the second mention of any word. Given that both emphatic stress and repetition are typical of infant-directed speech, it is important to document the pattern of words that caretakers produce (and the form of stress those words are produced with) beyond the first two mentions. It is also important to examine how this pattern might influence infants’ word learning.

Current Studies

The present studies were designed to determine, first, how mothers produce and naturally stress words across multiple mentions when addressing their infants and, second, how familiarization with words following different patterns of emphatic/nonemphatic stress might influence infants' subsequent recognition of those words in fluent speech. We recorded mothers narrating a simple puppet show to their infants and analyzed their spontaneously repeated words. We examined repetitions of particular words and observed that mothers methodically alternate between using emphatic and nonemphatic stress across successive mentions. We then systematically varied emphatic/nonemphatic stress on target words during both familiarization and recognition testing to determine how such stress influences 7.5-month-old infants' ability to detect words in fluent speech. The key questions we asked concerned how alternating degrees of acoustic emphasis affects infants' word recognition, whether infants' recognition is disrupted by changes in acoustic emphasis, and whether emphatic stress on newly introduced words facilitates their later recognition. Finally, we considered how mothers' productions and infants' perceptual predilections might dovetail.

Study 1

This study explored the patterns of emphatic stress that mothers display in their speech to infants. We used an elicitation method based on earlier work (Fisher & Tokura, 1995), which allowed us to influence the content of spontaneous speech that mothers directed to their infants.

Method

Participants—Speakers were 12 English-speaking mothers with infants between the ages of 9- and 10-months ($M = 280$ days; range = 269 days to 299 days).

Stimulus events—Mothers watched a puppet show with their infants seated on their laps and were asked to describe the simple events occurring in the show to their infants. A series of scenes were designed for mothers to view with their infants. In each, mothers were cued to produce specific content words in their interactions with the infants. All of the scenes had a common agent (a turtle puppet), but the scenes had different patients and actions. Before each scene, mothers saw a cue card with the names of the patient and action for that scene; mothers were instructed to explain the scene taking place in the puppet show to their infants using the noun and verb provided. For example, mothers would be cued with the combination “gazelle/push” prior to seeing the turtle begin to push the gazelle. They would then describe this scene to their infants, using those terms.

The puppet shows contained scenes in which actions were performed on eight different animals. These animals were chosen to have bisyllabic names. To control for possible influences of lexical stress (Echols, Crowhurst, & Childers, 1997) and to avoid artifacts attributable to word choice (see Vroomen, Tuomainen & de Gelder, 1998 for discussion), we selected target names such that half carried word-initial stress (e.g., stress on the first syllable of a two-syllable word) and half carried non-word-initial stress (e.g., stress on the second syllable of a two-syllable word). The resulting target words are listed in Table 1. The eight actions and eight puppets were presented twice across two blocks, resulting in a total of 16 scenes. The pairing of actions with animals was counterbalanced across the two blocks. Within each block, action-animal pairings were presented in a random order.

Scene presentation—The puppets were manipulated by an experimenter hidden behind a puppet stage. In an initial brief episode, the common puppet appeared alone. Each subsequent scene followed the same sequence: both puppets appeared together on the stage and remained still until the mother was shown the patient/action cue card. Then the experimenter began to

enact the scene repeatedly, while the mother described it to her infant. Each scene concluded when the mother stopped talking and remained silent for at least two seconds. The next scene was then shown in the same manner. By describing each scene in this way, mothers were encouraged to speak naturally to their infant, in a manner typical of infant-directed speech. Mothers' narratives of the puppet show were audio-taped and each recording was transcribed for subsequent analysis.

Data coding—Each mother's interaction with her infant was digitally recorded and transcribed in its entirety, with the onset of each new event clearly labeled within the transcript. Within each event sequence, every mention of a target word was highlighted. This was first verified via an acoustic comparison of the transcription to the corresponding recording. These highlighted transcripts then were used to guide localization of each mention of a target word for subsequent acoustic analysis. Acoustic analyses were carried out using the Bliss Speech Analysis System (developed at Brown University) and the Praat program (Boersma & Weenink, 2002). The primary acoustic correlates of focal stress noted by Fisher and Tokura (1995) were duration, minimum and maximum fundamental frequency (F_0), average F_0 , and overall range of F_0 . Each measurement averages across the entire word. These measures form the basis of the acoustic analyses of the speech stimuli collected here.

Results

Descriptive Analysis—Across the 192 scenes recorded, mothers produced a total of 669 tokens, averaging 55.75 (SD = 10.92) per person. This resulted in an average of 83.63 (SD = 9.86) tokens per item (animal name) across the twelve mothers whose speech was analyzed. In 3% of all scenes (6 scenes), mothers produced only a single mention of the target name. Most of these scenes occurred at the end of the recording sessions, when mothers and infants had clearly tired of the task. In 14% (26 scenes), mothers produced two mentions of the target. Twenty-eight percent of the scenes (54 scenes) elicited three mentions of the target; 31% (59 scenes) elicited four mentions, 17% (32 scenes) elicited five mentions, and 7% (12 scenes) elicited six or more mentions of the target. In one scene, a mother produced 10 mentions of the target!

Data on the total number of times target words were mentioned by individual mothers is shown graphically in Figure 1. Half of the mothers always repeated the target name beyond the first mention; Mothers 8 and 12 always repeated the target names at least three times. For most individuals, the modal number of mentions was three or four; all mothers mentioned at least one word four times, and all but one mother mentioned at least one word six or more times.

Acoustic Analyses—A summary of the acoustic analyses appears in Table 2. In keeping with Fisher and Tokura's analyses, we first compared our acoustic measures for the first and second mention of each target. The five acoustic measures for the first and second mentions of target words were each analyzed using a single-factor (mention) repeated measures ANOVA. All analyses were conducted with subjects as the random factor. These revealed that second uses of target words were significantly shorter than the first, $F(1, 11) = 26.64, p < .001, \eta^2 = 0.71$. Average F_0 was higher for first than second mentions of target words, $F(1, 11) = 4.91, p < .05, \eta^2 = 0.31$. Although the difference between minimum F_0 for the first and second mentions did not reach significance, $F(1, 11) = 4.25, p = .064, \eta^2 = 0.28$, the maximum F_0 did differ significantly between the two, $F(1, 11) = 21.74, p < .001, \eta^2 = 0.66$, as did the overall range of F_0 , $F(1, 11) = 44.44, p < .001, \eta^2 = 0.80$.

Overall, these findings are consistent with those reported by Fisher and Tokura across the first and second mention; they found comparable significant differences for all five measures. Thus, our data from comparing first and second mentions of target words comport with Fisher and

Tokura's claim that emphatic stress is reduced from first to second mention in speech directed to infants, as it is in speech between adults. However, we have many more subsequent mentions of each word within scenes to consider. What happens in the acoustic form of a word beyond the second mention? One interpretation would predict a monotonic decrease across repetitions in measures correlated with stress, a function of increasing levels of familiarity, of "givenness," across mentions. Statistically, this means that there should be a linear trend demonstrating a steady decrease in acoustic prominence across mentions: shorter duration, lower average and maximum F_0 , smaller F_0 range. Another interpretation, one in which stress is all or none, would not predict a linear trend, but rather something more like a big decrease in stress after the first mention followed by a plateau. A third would predict that caretakers revert to emphatically stressing the word again, were they to repeat it yet a third (and fourth) time, thus repeating the emphatic/nonemphatic cycle. To test these predictions, we compared changes in duration and acoustic indicators of stress across mention, focusing on the first through fifth mentions. We chose five mentions for analysis due to the limited number of tokens that speakers produced beyond five (21 tokens in all were produced as a sixth or subsequent mention); tokens from the first through fifth mentions provided us with 648 tokens for further analysis. Analysis of these data showed a marginally significant reduction of average duration across mention, linear trend, $F(1, 11) = 4.78, p = .051, \eta^2 = 0.30$. However, analyses of all the other acoustic measures revealed no other significant linear trends: average F_0 , $F(1, 11) = .232, ns$; maximum F_0 , $F(1, 11) = .001, ns$; minimum F_0 , $F(1, 11) = 1.32, ns$; and pitch range (reported here in semitones), $F(1, 11) = 2.45, ns$.

A strict view of the given/new contract would also predict the absence of higher-order trends. That is, a linear reduction in stress across mentions prior to pronominalization should preclude any other pattern of stress in the data. Nevertheless, our analyses revealed significant *quartic* trends in four out of the five dependent measures: rather than monotonically decreasing stress, mothers alternated their use of stressed and unstressed word forms. For example, although the average duration across the bisyllabic target words was 720 msec ($SD = 270$), a trend analysis across mentions by mothers revealed that the average word duration alternated significantly from mention to mention, quartic trend, $F(1, 11) = 17.42, p < .01, \eta^2 = 0.62$. A closer examination of these measures indicated that speakers produced words with relatively longer durations upon first mention ($M = 842$ msec, $SD = 294$), followed by a significant decrease in duration for the subsequent (second) mention ($M = 652, SD = 235$). However, rather than continuing subsequent repetitions with relatively reduced durations or the same duration as the second mention (as might be the case if the trend were strictly linear) mothers again produced words with relatively longer durations upon third mention ($M = 712$ msec, $SD = 279$) and continued alternating between longer and shorter production in this manner up through the fifth mention. This back-and-forth process can be seen in Figure 2, in which duration changes across mentions are plotted. The marginally significant linear trend can be observed embedded in this longer/shorter pattern, such that overall duration decreases with each subsequent mention. But the quartic trend is predominant.

A within-subjects trend analysis of maximum F_0 revealed a significant alternation from mention to mention as well, quartic trend, $F(1, 11) = 22.72, p < .001, \eta^2 = 0.67$. This alternating pattern is further reflected in a significant quartic trend across mentions for average F_0 , $F(1, 11) = 9.94, p < .01, \eta^2 = 0.47$. Finally, pitch range (reported here in semitones) followed the same repeating, alternating pattern (e.g., large range followed by small range) when analyzed across mentions, quartic trend, $F(1, 11) = 25.85, p < .001, \eta^2 = 0.70$. Of all the acoustic indicators of stress, only the average minimum F_0 did not significantly alternate from mention to mention in lockstep with the other acoustic indicators of stress, quartic trend, $F(1, 11) = .368, ns$.

Discussion

Sometimes, mothers say the same thing over and over and over again. Although we did not instruct mothers to repeat any words, they nevertheless did so readily. In only a tiny fraction of the scenes presented – mostly at the end of the session when mother and infant were clearly fatigued – did mothers fail to mention target names more than once. On average, mothers mentioned the target names 3.5 times per scene. Our acoustic analyses showed that first mentions are longer, higher pitched, and have greater pitch range than do second mentions. These results comport with those of Fisher and Tokura (1996), who suggested that something like a given/new contract is honored in child-directed speech.

Findings from Study 1, however, are not predicted by traditional models of emphatic stress as it is used in ADS (e.g., Chafe, 1976; Clark & Haviland, 1977; Halliday, 1967; Haviland & Clark, 1974, Prince, 1981). These predict that words should receive full emphasis when first introduced and reduced stress with subsequent mentions (either gradually or completely). As the significant quartic trends that we found show, these data did not fit this pattern, but neither did they fit patterns one would expect if one simply predicts that emphatic stress aids infants' language learning (e.g., Bernstein & Ratner, 1996). Although mothers in Study 1 repeated words in a manner characteristic of infant-directed speech, they did not consistently reduce or emphasize stress on target names across mentions, nor did they shift to using pronouns. Rather, on average, they alternated between producing the full form of the target word with emphatic and nonemphatic while speaking to infants, regardless of how many times they mentioned the given word.

Of course, we should expect some variability in stress-related acoustic features unrelated to its informational status, since it is only one of a number of influences on the acoustic properties of words. Other factors include the position of the word within an utterance and utterance length, to name just two. Since the alternating pattern we have observed is the result of averaging across talkers and words, it undoubtedly is influenced by these other factors as well, and the given/new characterization of the pattern is one of many possible such characterizations. What is important to note, however, is that this pattern appears to be specific to infant-directed speech and is consistent with our third prediction: That is, it is the product of multiple, sequential mentions of particular words, a phenomenon not readily observable in speech directed towards adults. Overall then, our findings are consistent with the prediction that caretakers revert to the emphatic/nonemphatic stress pattern for subsequent mentions of a word, apparently repeating that cycle as many times as their repetition of the word required.

It is important to keep in mind, however, that the nature of the task required mothers to engage infants' attention given a repeated visual scene. Perhaps caretakers cycled between emphatic/nonemphatic stress in order to maintain their infants' attention to the extended scene. One can imagine that the acoustic result of such cycling would be a rhythmic, almost sing-song pattern that may have helped maintain infants' attention. Another possibility is that caretakers were producing more complex patterns of stress across their utterances, which our focus on target animal names failed to capture. Although we did not formally code the form of stress used to refer to the other animal (i.e., the turtle), one can imagine that caretakers may have shifted the focal stress from one acting animal to the other across mentions. Regardless, caretakers' repetition of the emphatic/nonemphatic stress pattern across mentions of same word provided infants with varied acoustic examples of targets in a manner that might aid in later recognition (see Singh, 2008). To determine whether infants can indeed recognize words as they follow this alternating stress pattern within fluent speech, Study 2 extended the naturally occurring pattern of emphatic stress observed in Study 1 to an infant-word recognition paradigm.

Study 2

In this set of three experiments, we explore the role of emphatic stress in infants' spoken word recognition using the word recognition technique—the headturn preference procedure, or HPP—described earlier. To date, studies investigating infants' spoken word recognition have not explicitly manipulated emphatic stress of target words. We presume, however, that because familiarization exemplars have been produced explicitly for the experiments in which they were used, it is likely that they have been uniformly emphatically stressed. Therefore, in the following three experiments we manipulated the pattern of emphatic/nonemphatic stress on target words that infants heard during the familiarization and recognition phases of each.

Study 2A

Results from Study 1 indicate that mothers naturally alternate between emphasizing and de-emphasizing content words in their speech to infants. To test whether infants are sensitive to this form of acoustic variation, in Study 2A we familiarized infants with words following this alternating stress pattern and subsequently tested whether they recognized the words they had been familiarized with in isolation when they were in fluent speech but followed the same alternating stress pattern. In each of the subsequent perceptual studies, we familiarized infants with isolated words and tested them with those words in sentences, as this design allowed us to establish infant recognition for words in fluent speech. We did this rather than familiarizing the infants with words in fluent speech and then testing them on those words in isolation, as this is the first set of studies to explicitly examine the influence of emphatic/nonemphatic stress on infant word recognition. The pattern of testing we employed is generally considered an easier task for infants and thereby will serve to establish whether manipulations of emphatic/nonemphatic stress patterns on familiarization items influences subsequent infants' subsequent recognition of those items. Future research will extend testing to familiarization with words in fluent speech and recognition testing with isolated words.

Method

Participants: Sixteen infants (7 females and 9 males) from monolingual English-speaking households were tested. The average age of the infants was 230 days ($SD = 5$ days; range = 220 to 240 days), approximately 7.5 months. Testing was not completed with two additional infants due to crying and fussiness, and with one due to equipment failure.

Stimuli and apparatus: Stimuli consisted of four bisyllabic words and four six-sentence sets. The words were animal names (*chicken*, *dolphin*, *falcon*, and *monkey*) that had been judged unfamiliar to infants of this age in pretesting with mothers. We used only strong-weak target words because, as noted earlier, research has indicated that 7.5-month-olds can recognize strong-weak, but not weak-strong, bisyllables in fluent speech (Jusczyk, Houston, & Newsome, 1999). These stimuli were embedded in fluent-speech utterances that were designed to mimic the alternating emphasis patterns generated naturally by mothers in Study 1.

Briefly, we established that the acoustic characteristics of our stimuli were consistent both with those reported in the literature for emphatic relative to nonemphatic stress (Fisher & Tokura, 1995; Fowler, 1988; Fowler et al., 1997), and with the values reported in Study 1. Similar to previous studies and our own earlier observations, emphatically stressed words were characterized by higher overall F_0 maxima and higher mean pitch and pitch range than nonemphatic stress. Although speech rate did not differ significantly for the two types of sentences, duration of target words within the sentences did differ by focal stress type. Based on independently calibrated perceptual ratings and acoustic analyses, the stimuli were judged to convey the appropriate forms of emphatic and nonemphatic stress as produced by mothers addressing their infants. With these sentences and single word tokens, we were able to construct

the familiarization and test stimuli for our perceptual experiments controlling for speaker variation. Appendix A includes more detailed information on how utterances were developed for this and the subsequent studies; Appendix B shows the sentences used in this study.

Testing Apparatus: Testing was conducted in a three-sided booth constructed of pegboard that was placed inside a sound-treated laboratory room. Each of the three walls of the booth was 120 cm wide. A chair was placed at the open side of the testing booth for the parent to sit on with the infant in his or her lap, about 110 cm from the center wall of the booth. A single, amber light was mounted at an infant's eye level (86 cm above the floor) on the booth's center wall. Single green lights were mounted on each of the two side walls at the same level as the center light. Loudspeakers were positioned behind the side walls, below the two green lights. A video camera (Panasonic CCTV model VW-1410) was situated behind the center wall with its lens trained through a hole cut 12.3 cm above the yellow light. Only the lens was visible from within the testing booth. The camera's view encompassed the width of the testing booth and allowed infants' behavior to be remotely monitored. The loudspeakers and lights were linked to a computer in a control room located down the hall from the test room. A video recorder (Panasonic Time Lapse model AG 6040) and monitor (Panasonic model WV-5410) were connected to the video camera in the test room so that an experimenter seated in the control room could observe and record the infants' responses, while at the same time not hearing what the infants were hearing. The experimenter also controlled the onset and offset of the three lights and the presentation of auditory stimuli to the infant via a computer program designed specifically for operating the experiment. All stimuli were set to play at a conversational volume (75 dB) in the testing room using a Radio Shack sound level meter.

Testing Protocol: Infants heard repetitions of two different target words during familiarization. Presentation of the two words was randomized across trials such that during any given trial, only one target word was presented, alternating between being emphatically and nonemphatically stressed production. The pairs of words used as familiarized targets (either *chicken* and *dolphin* or *monkey* and *falcon*) were counterbalanced across subjects. During recognition testing, infants heard sentences containing each of the four words, so that the familiar words in sentences for some infants were the unfamiliar words to others, and vice versa. All the sets of test sentences were arranged so that acoustic production of the target word within each alternated between emphatic and nonemphatic stress from one sentence to the next within a set.

During testing, the infant was seated on a parent's lap facing the center light. The parent listened to instrumental music over Bose aircraft-quality noise-cancellation headphones to mask experimental stimuli. Each trial began with the amber light on the center panel blinking in order to draw the infant's attention to the center of the booth (to midline). When the infant oriented toward the blinking center light, the experimenter called for a trial. At this point, the center light was extinguished and one of the green lights located on either side of the infant began to blink. Side of presentation was randomized across trials. When the infant oriented in the direction of the blinking green light, the experimenter in the control room pushed a button that caused the blinking light to illuminate and the auditory stimuli to begin playing through the loudspeaker on that side of the testing booth. The change in the light occurred simultaneously with the onset of the auditory stimuli to help infants form an association between their own orientation towards the light with the onset of auditory stimuli.

During the familiarization phase of the experiment, the target words continued playing until the infant turned away, up to a maximum of 30 seconds. A trial automatically terminated if the infant looked away for more than 2 seconds and a new trial began. If the infant turned briefly away from the target, but for less than 2 seconds, the trial continued with time spent looking away not included in the calculation of total orientation time for that trial. Familiarization

continued until the infant received 30 seconds of exposure to each of the two target words. Importantly, once the infant achieved the familiarization criterion for one word, the familiarization trials that followed presented only the other word, until criterion was reached for the second word as well. This modification of the HPP was instituted to ensure that differences in orientation times during recognition testing could not be due to different amounts of familiarization with the two target words. In each experiment, familiarization stimuli were counterbalanced across subjects. When the infant reached 30 seconds of looking time with the second word, the test phase began.

During the recognition test phase, all of the infants heard four sets of concatenated sentences (see Appendix B). Two of the sets contained sentences with familiarized words and two contained sentences with non-familiarized words. Recognition test trials were blocked so that each of the four sentence sets occurred once within a given block. A total of three blocks were presented to each infant and the order of sentence sets within a block was randomized. Within each set, the order of pairs of sentences—where each pair contained one sentence with emphatic stress on the target word followed by one sentence with nonemphatic stress on the target word—was randomized on each trial. The recognition test procedure was identical to the familiarization procedure, except that the side light continued to blink while the infant was oriented towards it. As in the familiarization phase, a trial automatically terminated if the infant looked away for more than 2 seconds; the looking time for that trial was recorded based on the point at which the infant looked away. If the infant continued to look at the light for 20 seconds, the trial ended automatically and the next trial began. If the infant failed to look at the light for at least 2 seconds, the trial repeated automatically, with a new randomized order of the sentence pairs for that set. A minimum criterion of 2 seconds was necessary for the infant to hear at least one instance of the familiar or unfamiliar word in a single sentence.

Results and Discussion

Analysis of Familiarization Phase: A within-subjects analysis of familiarization showed no difference in the number of trials infants received with one target word ($M = 3.75$, $SD = 1.06$) versus the other ($M = 4.0$, $SD = 1.33$), $t(15) = 0.49$, ns. Nor was there any difference in the amount of orientation per familiarization trial for one target ($M = 8.54$ s, $SD = 2.80$) versus the other ($M = 8.30$ s, $SD = 4.54$), $t(15) = 0.74$, ns. Therefore, infants did not receive different amounts of familiarization to the two target words, nor did they complete familiarization at different rates for the target words.

Analysis of Recognition Test Phase: The dependent measure for recognition was the time an infant spent oriented towards the light while the different sentence sets played. Improved recognition as a result of familiarization was operationalized as a significant difference in orientation towards those sentences containing familiarized words, relative to those containing unfamiliar words. Mean orientation times to the four different sentence sets were thus calculated for each infant across the three test blocks. These were then averaged for sentence sets containing the familiar words and for those containing the unfamiliar words. These orientation times are presented in Figure 3. Overall, infants oriented to sentence sets containing familiar words for longer periods ($M = 8.29$ s, $SD = 2.61$ s) than they did for those containing unfamiliar words ($M = 6.70$ s, $SD = 1.54$ s), $t(15) = 3.01$, $p < .01$. A mixed-factor analysis of variance further indicated that this effect was not influenced by counter-balancing condition, $F < 1$. Effects of counter-balancing were examined in all studies and no interactions with key variables were found; in all subsequent analyses, counterbalancing conditions were collapsed.

Study 2A demonstrated that infants can recognize words in fluent speech with which they have been previously familiarized when those words alternate between emphatic and nonemphatic stress in a manner consistent with the natural production observed in Study 1. What remains

unclear is whether familiarization with a mix of stressed and unstressed tokens results in *better* word recognition than familiarization with words that receive only emphatic stress. Studies 2B and 2C were designed to address this question.

Study 2B

In Study 2B, two groups of 7.5-month-old infants were tested in a manner identical to that described in Study 2A. However, rather than being familiarized with words whose stress alternated between emphatic and nonemphatic and then being tested with the same alternating pattern, half of the infants in this experiment were familiarized with items produced entirely with emphatic stress and half were familiarized with items produced entirely with nonemphatic stress. All infants were tested using the same sentence sets that were used in Study 2A, in which the stress on target and control words regularly alternated between emphatic and nonemphatic forms. This means that, for both groups of infants, familiarization items were acoustically similar to test items half of the time (and acoustically dissimilar for the other half). To the extent that emphatic stress aids word-form learning (e.g., Bernstein-Ratner, 1986), we expected that infants who were familiarized with words produced entirely with emphatic stress would be more effective at subsequently recognizing those words in fluent speech than infants who were familiarized with words produced with nonemphatic stress.

Method

Participants: Thirty-two infants (15 females and 17 males) from monolingual English-speaking households were tested. The infants' average age was 227 days ($SD = 5$ days; range = 218 to 243 days), approximately 7.5 months. Three additional infants were excluded due to general inattention and failure to complete the test.

Stimuli, apparatus, and procedure: One group of infants was familiarized with two target words produced with emphatic stress. A second group of infants was familiarized with two targets produced with nonemphatic stress. All infants were tested with four sentence sets in which target and control words alternated between emphatic and nonemphatic stress, as in Study 2A. All other aspects of the apparatus and procedure were identical to the previous experiment as well.

Results

Analysis of Familiarization Phase: Although it is important for infants to receive comparable amounts of familiarization with both targets, one might predict that they would complete familiarization earlier (that is, over fewer trials) for the words produced with emphatic stress than for the words produced with nonemphatic stress, given the acoustic salience of the former. Indeed, infants took slightly fewer trials to reach familiarization criterion with emphatically stressed words ($M = 3.94$, $SD = 0.64$) relative to nonemphatically stressed words, ($M = 4.25$, $SD = 1.74$), but this difference was not significant $t(31) = 0.11$, ns. Furthermore, individual trials for familiarization with emphatically stressed words lasted somewhat longer than those for nonemphatically stressed words ($M = 8.17$ s, $SD = 1.53$ vs. $M = 7.93$ s, $SD = 3.22$), however, this difference was only marginally significant, $t(31) = 0.54$, ns.

Analysis of Recognition Test Phase: The results of Study 2B can be seen in Figure 4. Data were analyzed in a 2 (stress type) \times 2 (familiarity) mixed ANOVA. During recognition trials infants oriented longer to the words with which they had been familiarized (7.96 s; $SD = 2.08$ s) than to unfamiliar words (6.83 s; $SD = 1.80$ s), $F(1, 30) = 12.53$, $p < .001$, $\eta^2 = 0.08$. There was also a main effect for stress type: infants who were familiarized with emphatically stressed words produced longer overall orientation times (averaged across target and control sentence sets) at recognition (8.18 s; $SD = 2.02$ s) than did infants who were familiarized with words

carrying nonemphatic stress (6.61 s; SD = 1.69 s), $F(1, 30) = 8.20, p < .008, \eta^2 = 0.16$. However, orientation times towards sentence sets containing familiar words did not differ significantly based on whether that familiarization took place with emphatic or nonemphatic stress, as there was no interaction between condition and familiarity, $F(1, 30) = .054, ns$.

Discussion—First, to establish the relative strength of the effects observed in this and the previous study, we computed the effect sizes for each by calculating the pooled standard deviation from the original standard deviations rather than from the t statistic, and constitutes a more conservative approach (Dunlop, Cortina, Vaslow, & Burke, 1996). This yielded an effect for Study 2A that was moderately strong (Cohen's $d = 0.74$). Because we are interested in the arm of Study 2B in which infants were familiarized with emphatically stressed words, we calculated the effect size specifically for it (Cohen's $d = 0.53$). The moderate effect size in Study 2B is notably smaller than that obtained in Study 2A, when stress at familiarization contained both emphatic and nonemphatic forms. This suggests two possibilities: that variability in the acoustic structure of the target words during familiarization is at least as important as the acoustic salience of the words themselves, or that similarity in the acoustic structure of target words from familiarization to recognition bolsters the effect, regardless of what that structure is. We will return to this issue in Study 3.

Second, even when familiarization stimuli were acoustically consistent only half of the time with items at recognition, infants were still able to recognize familiar targets. If acoustically salient, emphatic stress at familiarization were driving infants' subsequent recognition in this more challenging task, then infants familiarized with acoustically reduced items should have had difficulty recognizing them. Instead, infants familiarized entirely with nonemphatically stressed stimuli were able to recognize alternating stress target items at recognition to the same degree as infants familiarized with emphatically stressed stimuli.

Finally, although the main effect of stress type in Study 2B indicates that infants' attention was affected by the nature of the familiarization stimuli, we (somewhat unexpectedly) did not observe that emphatically stressed words were immediately superior in engaging infants' attention, as there were no differences in orientation time during the familiarization phase of the experiment. Rather, emphatic stress during familiarization had the effect of better *maintaining* infants' attention during the recognition phase of the procedure. Insofar as mothers' immediate goal is to maintain their infants' attention in ongoing interactions, this might explain why mothers revert to emphatic stress when repeating words. Nevertheless, a more sensitive design might yet reveal effects of emphatic stress on infants' word-form learning *for* recognition. We thus repeated Study 2B, but used a within- rather than between-subjects design.

Study 2C

The experiments reported up to this point have involved manipulations in which infants were familiarized with words produced with a single pattern of stress (all emphatic, all nonemphatic, or alternating between the two). Does acoustic salience assume additional importance when emphatic and nonemphatic stimuli are pitted against one another? In Study 2C, one of the words that infants heard was always emphatically stressed during familiarization, while the other was always nonemphatically stressed. Target words in sentence sets, as in Studies 2A and 2B, alternated between emphatic and nonemphatic stress. If acoustic similarity is the only factor determining infants' success at word-form recognition, then we should expect equally strong recognition of both familiarization items. Alternatively, if emphatic stress does facilitate word-form learning, then we should observe stronger recognition of the word that was emphatically stressed during familiarization.

Method

Participants: Twenty-four infants (10 females and 14 males) from monolingual English-speaking households were tested. The infants' average age was 231 days ($SD = 8$ days; range = 211 to 242 days), approximately 7.5 months. Three additional infants were excluded due to inattention and failure to complete the test.

Stimuli, apparatus, and procedure: Infants were each familiarized with two words, one produced with emphatic stress and the other produced with nonemphatic stress. As in the previous experiments, half of the infants heard *chicken* and *dolphin* during familiarization, whereas the others heard *monkey* and *falcon*. Within each of these subgroups, assignment of words to emphatic and nonemphatic stress was counterbalanced. The recognition test phase used the same sentence sets as those used in Studies 2A and 2B, with stress on target and control words alternating between emphatic and nonemphatic. The apparatus and procedure were the same as in the previous experiments, except that type of stress at familiarization was manipulated as a within-subjects variable.

Results

Analysis of Familiarization Phase: Infants in the two conditions did not receive different amounts of exposure to the different target words during familiarization, nor did they complete familiarization at different rates for the two target words, as measured by mean looking time per trial. A within-subjects analysis showed no difference in the number of familiarization trials infants received with emphatically stressed words ($M = 5.04$, $SD = 1.95$) versus nonemphatically stressed words ($M = 5.42$, $SD = 1.47$), $t(23) = 0.16$, ns. Nor was there any difference in the amount of orientation time per familiarization trial for emphatically stressed words ($M = 6.55$ s, $SD = 3.13$) versus nonemphatically stressed words ($M = 6.02$ s, $SD = 2.79$), $t(23) = 0.08$, ns.

Analysis of Recognition Test Phase: Average orientation times were analyzed in a one-way repeated measures (following emphatic familiarization, following nonemphatic familiarization, no familiarization) ANOVA and can be seen in Figure 5. Overall, mean orientation times for the three conditions were significantly different from one another, $F(2, 46) = 23.04$, $p < .001$, $\eta^2 = 0.18$. As expected, orientation times towards sentence sets containing words that were familiarized with emphatic stress ($M = 7.87$ s; $SD = 3.06$ s) were significantly longer than orientation times towards sentences sets containing unfamiliar words ($M = 5.04$ s; $SD = 1.83$ s), $t(23) = 5.80$, $p < .001$, Cohen's $d = 1.12$. Orientation times to sentence sets containing words that were familiarized with nonemphatic stress ($M = 6.09$ s; $SD = 2.60$ s) were also significantly longer than orientation times to sentence sets containing nonfamiliarized words, $t(23) = 3.64$, $p < .001$, Cohen's $d = 0.47$. These results parallel the influence of familiarity observed in Study 2B.

Of greater interest, infants oriented significantly longer to sentence sets containing words that they had initially heard produced with emphatic stress than those containing words initially heard produced with nonemphatic stress, $t(23) = 3.90$, $p < .01$. The difference in orientation times observed here could reflect the superior attention-attracting and attention-maintaining qualities of words stressed emphatically at familiarization, or the overall preference an infant has for emphatically stressed words. Either interpretation is consistent with mothers' repeated use of emphatic stress across of mentions of words to their infants. Yet despite the more sensitive, within-subjects design we employed here, the effect size for the result, Cohen's $d = 0.63$, is still somewhat smaller than that observed in Study 2A, when stress at familiarization precisely matched stress at recognition. Study 3 was thus designed to address the interactive influence of emphatic stress and matching stress on early infant word learning and recognition.

Study 3

Several studies have found effects of similarity across a word's acoustic form in early infant word recognition, but these have focused on different acoustic characteristics than the present investigation. For instance, Houston and Jusczyk (2000) familiarized 7.5-month-olds with stimuli produced by a female talker and found that they listened longer to test passages containing the familiarized words if the sentences were also produced by a female talker, rather than a male talker. Similarly, Singh, Morgan, and White (2004) found that 7.5-month-olds recognized words when speaker affect matched, but not when speaker affect varied, across familiarization and testing. Singh, White, and Morgan (2008) digitally raised and lowered the pitch of both words and sentences and found that 7.5-month-olds recognized words only when familiarization and test pitch matched. However, infants do not incorporate all correlated vocal properties into their word-form representations: Singh et al. (2008) showed that word form recognition is unimpeded when amplitude is systematically varied across familiarization and recognition testing.

Do young infants treat emphatic stress more like talker gender (Houston & Jusczyk, 2000), speaker affect (Singh et al., 2004), and speaker pitch (Singh et al., 2008), or instead like amplitude (Singh et al., 2008)? In Study 1, we showed that mothers systematically vary emphatic stress across repeated mentions of words, a practice that should, according to the view developed here, lead to infants' exclusion of emphatic stress from their word-form representations. When this might occur, however, is unknown. To our knowledge, all of the examples of word repetition in the literature are drawn from conversations with infants who have already begun to speak; such repetition may be particularly likely to occur when caretakers are consciously trying to teach words to their infants, and whether caretakers repeat words to preverbal infants is an open question.

In this study, we orthogonally manipulated the presence of emphatic or nonemphatic stress during the familiarization and recognition phases of the experiment to determine their effects on word recognition in a between-subjects design. The four resulting conditions allowed us to address each of four possibilities: (1) that emphatic stress is facilitative to word recognition when present during familiarization, (2) that emphatic stress is facilitative to word recognition when present at recognition, (3) that emphatic stress must be present at *both* familiarization and recognition to be facilitative, or (4) that emphatic stress is not particularly facilitative to word recognition.

Method

Participants—Eighty infants (34 females and 46 males) from monolingual English-speaking households in were tested. The infants' average age was 230 days ($SD = 8$ days; range = 213 to 246 days), approximately 7.5 months. Eight additional infants were excluded due to inattention and failure to complete the test.

Stimuli, apparatus, and procedure—Four groups of twenty infants were tested in each of the four conditions. Stress type at familiarization and at recognition were orthogonally manipulated as between-subject independent factors, each with two levels (emphatic stress or nonemphatic stress). Infants in each condition were familiarized with two English target words, produced either with emphatic or nonemphatic stress. They were then tested for recognition, again using words produced either with emphatic or nonemphatic stress. The apparatus and test procedure were identical to those in the previous experiments; as described in Appendix A, additional stimuli were recorded and rated to round out the sentence sets.

Results

Analysis of Familiarization Phase—Infants in the four conditions did not receive different amounts of familiarization to emphatically and nonemphatically stressed words. Analysis of the average number of trials among the four familiarization conditions revealed no significant difference, $F(1, 76) = 1.40$, ns, nor was there any difference in the average length of each trial, $F(1, 76) = 1.63$, ns. Specific comparisons between familiarization conditions revealed no significant difference for familiarization with emphatic stress relative to nonemphatic stress either in average orientation time per trial ($M = 7.85$ s, $SD = 1.64$ vs. $M = 7.27$ s, $SD = 1.55$), $t(78) = 1.35$, ns, nor a significant difference in average number of trials ($M = 4.16$ s, $SD = .96$ vs. $M = 4.43$ s, $SD = 1.04$), $t(78) = -1.17$, ns.

Analysis of Recognition Test Phase—In each of the four conditions, average orientation times were calculated for sentence sets containing familiarized and nonfamiliarized words. Data were entered into a 2 (Stress at Familiarization: Emphatic vs. Nonemphatic) \times 2 (Stress at Recognition: Emphatic vs. Nonemphatic) \times 2 (Familiarity: Familiar vs. Novel) mixed ANOVA, with Familiarity as the only within-subject factor. Overall, there was a main effect of Familiarity, $F(1, 76) = 5.34$, $p < .05$, $\eta^2 = 0.06$, such that infants spent more time oriented towards sentence sets containing familiar words ($M = 6.9$ s, $SD = 1.946$) than unfamiliar words ($M = 6.4$, $SD = 1.946$). However, contrary to research arguing for the importance of emphatic stress (e.g., Bernstein Ratner, 1996), there was no significant effect of stress at familiarization, $F < 1$. Only the main effect of stress at recognition approached statistical significance, $F(1, 76) = 3.22$, $p < .08$. More telling, the only other statistically significant effect was a three-way interaction between stress at familiarization, stress at recognition, and familiarization, $F(1, 76) = 7.46$, $p < .05$, $\eta^2 = 0.08$. The means and standard deviations for this interaction are shown in Table 3.

These results reveal the importance of acoustic similarity—or stress matching—between familiarization and recognition. We can consider the evidence in two ways. First, recognition (operationalized as significant differences in orientation time towards sentence sets containing familiarized versus nonfamiliarized words) occurred when stress was matched across familiarization and recognition, but not when it was mismatched. When target words occurred with emphatic stress at both familiarization and recognition, infants oriented significantly longer to sentence sets containing familiarized words, $t(19) = 2.13$, $p < .05$, Cohen's $d = 0.52$. The same pattern appeared when target words were consistently produced with nonemphatic stress, $t(19) = 2.44$, $p < .03$, Cohen's $d = 0.48$. In contrast, when target words were produced with one type of stress at familiarization and the other at recognition, orientation times to sentence sets with familiarized versus nonfamiliarized words did not differ; in both cases, $t(19) < 1$.

Second, we analyzed the between-subject determinants of the recognition scores (computed by subtracting orientation times for sentence sets containing nonfamiliarized words from orientation times for sentence sets containing familiarized words). These are shown in Figure 6. When we grouped the data according to type of stress at familiarization or type of stress at recognition, there were no differences in recognition scores, $F_s < 1$. However, when we grouped the data according to matching versus mismatching stress, recognition scores were significantly higher in the matched conditions than in the mismatched conditions, $F(1, 78) = 7.46$, $p < .01$, $\eta^2 = 0.09$. These findings highlight the importance of acoustic similarity from one encounter with a word (e.g., during familiarization) to the next (e.g., during recognition) in supporting infants' emerging word recognition abilities.

Discussion

Study 3 was designed to explore why familiarization with alternating emphatic/nonemphatic stress in Study 2A produced a larger effect size than when familiarization took place entirely with emphatic stress in Study 2B. We had hypothesized that providing familiarization with both emphatic and nonemphatic forms of stress (rather than with emphatic forms only) might strengthen recognition by providing a better match with recognition targets that were sometimes emphatic and sometimes nonemphatic. The results of Study 3 indicate that infants' ability to recognize words in fluent speech following familiarization does depend, at least to some degree, on the similarity of stress present from familiarization to recognition. That is, for 7.5-month-old infants, emphatic stress was facilitative at familiarization when emphatic stress was also present at recognition. Similarly, nonemphatic stress was facilitative at recognition when nonemphatic stress was also present at recognition. Moreover, when there was an "emphasis mismatch" at familiarization and recognition (e.g., emphatic stress at familiarization and nonemphatic at recognition; nonemphatic stress at familiarization and emphatic at recognition), little to no recognition was observed.

These results comport with other studies that have shown matching effects in early infant word recognition. However, if what is critical for infants' early ability to recognize words is the acoustic similarity between familiarization and recognition (viz., earlier and later instances of words), then *any* pattern of usage that mothers might use should do equally well to support learning, provided that it is consistent. In that case, if mothers must modify the way they use emphatic/nonemphatic stress to accommodate immature listeners, the simplest way (at least theoretically) would be to monotonically reduce stress on words across mentions, albeit with a shallower slope than that found in adult-directed speech, thereby providing infants with the range of possibilities for that dimension of stress. Our finding of no significant linear reductions across mention for any of the pitch measures in Study 1, and only a marginally significant reduction in word duration across mention, clearly shows that this is not the pattern observed with mothers. Rather, mothers' behavior appeared to conform to the common-sense notion that the attention-getting nature of emphatic stress is important for infants' word learning and is a useful tool for framing the presentation of alternate versions of the word (e.g., nonemphatically stressed instances of the same word).

General Discussion

In this article, we report two related sets of findings, the first concerning prosodic properties of repeated words in infant-directed speech, and the second concerning effects of such properties on infants' spoken word recognition. From the inception of formal study of child-directed speech (Ferguson, 1964; Snow, 1972), researchers have noted the high frequency of exact and periphrastic repetitions of phrases and sentences; the individual words contained in these necessarily are repeated as well. Previous research by Fisher and Tokura (1995) examined the prosodic properties of the first and second mentions of repeated words in speech to 14- and 15-month-olds and to adults. Observing that second mentions were reduced (shorter duration, lower pitch, smaller pitch excursions) in both registers, Fisher and Tokura concluded that something akin to a given/new contract (Clark & Haviland, 1977; Fowler & Housum, 1987) is observed in child-directed speech in much the same way that it is in adult-directed speech. In the present work, we adopted Fisher and Tokura's procedure and induced mothers to spontaneously mention target words to their infants multiple times. Our analyses of first and second mentions comported with their earlier findings. However, our analyses of third, fourth, and fifth mentions produced unexpected results: rather than monotonically or categorically reducing prosodic properties across mentions, mothers generally oscillated between producing non-reduced, emphatic tokens and reduced, nonemphatic tokens, following a damped quartic trend (see Figure 2). These results do not strictly conform to the given/new model. Across

repeated mentions, mothers produce multiple emphatically stressed instances of the word, a phenomenon that is not observed in the suprasegmental cues in adult-directed speech. However, this is in large part because adult-directed speech does not include multiple, sequential repetitions.

A theme implicit in much research on language development has been that such repetition with emphasis is likely to aid word recognition by making words perceptually prominent (Aslin, Woodward, LaMendola, & Bever, 1996; Bernstein Ratner, 1996; Cooper & Aslin, 1990; Fernald, 2000; Fernald & Mazzie, 1991; Werker & McLeod, 1989). The results of Study 2C are consistent with this: when infants were tested with sentence sets in which target words appeared with alternating emphatic and nonemphatic stress, recognition was slightly enhanced for words that had first been familiarized with emphatically stressed rather than nonemphatically stress. That emphatic stress assumes additional prominence in a more complex situation should not be surprising. In everyday life, it is likely that the contrast provided by emphatic stress assists caretakers in focusing infants' attention on particular items, despite the welter of words in the caretaker's speech stream. By presenting both forms of the stress during familiarization in Study 2C, we may have drawn their attention to this dimension in a way that would not happen outside the laboratory, when it is unlikely that a caregiver would alternate between the presentations of two new words, while emphasizing only one. Future studies, in which a reversed design (e.g., familiarization with sentence sets and recognition testing with single words) is followed, will further elucidate the specific influence of emphatic stress on lexical acquisition.

In addition to finding at least some facilitation of recognition for emphatically stressed words, an advantage that appears to hold across the life span (cf. Goodman, Nusbaum, Lee, & Broihier, 1990), we also found evidence for what might be considered an infant-specific pattern of recognition. When infants were tested with sentence sets in which target words appeared with exclusively emphatic or nonemphatic stress, recognition was successful *only* when infants had first been familiarized with instances of the word bearing the same form of stress. In Study 3, 7.5-month-olds who were familiarized and tested with emphatically stressed words and sentence sets, or familiarized and tested with nonemphatically stressed words and sentence sets, showed recognition scores that were significantly greater than zero (see Figure 6). In contrast, infants who were familiarized with emphatically stressed words and tested with nonemphatically stressed sentence sets, or vice versa, showed recognition scores that were not different from zero.

Lexically Relevant and Irrelevant Forms of Variation in Early Word Recognition

The results of Study 3 are consistent with several other studies that have found that early spoken word recognition may be disrupted by variation along a number of dimensions that are, to more mature listeners, lexically irrelevant. Houston and Juszyk (2000) found that 7.5-month-olds familiarized with tokens from a female talker failed to recognize words in sentences from a male talker. Singh, Morgan, and White (2004) found that 7.5-month-olds familiarized with tokens produced in one affect failed to recognize words in sentences produced in a different affect, even though the talker remained constant. Singh, White, and Morgan (2008) similarly found that 7.5-month-olds familiarized with tokens produced with raised pitch failed to recognize words in sentences produced with lowered pitch, and vice versa. Collectively, these results indicate that young infants are forming detailed representations of input stimuli and that they are weighting lexically-relevant and -irrelevant dimensions more or less equally.

Although such unbiased weightings are suboptimal for word recognition in any given language, they provide an optimal starting point for phonological learning. For infants learning Finnish, for example, segment duration will prove to be lexically relevant, for infants learning Mandarin, pitch contour will be relevant, and for infants learning Tamil, the contrast between dental and

retroflex places of articulation will be relevant. For English-learning infants, none of these will be relevant, but there is no way for them to know this in advance. Initial unbiased weightings combined with attention to statistical properties of the input – the characteristics of speech sound distributions and covariance among properties of those sounds – ensure that infants will not overlook relevant dimensions of variation, while providing them with a means for adapting these weightings in an appropriate, language-specific manner.

One consequence of this account is that infants may be easily misled in early stages of acquisition. As Study 3 and the above-cited studies show, around the time that infants are beginning to recognize spoken words in fluent speech (around six or seven months), we can easily confuse them by artificially conflating particular sequences of phones (that are lexically relevant) with any of a host of lexically irrelevant properties of speech. As infants gain more knowledge of the words of their language and adapt their weightings, it should become progressively more difficult to “fool” them. Indeed, in the instances of paralinguistic properties of talker identity, speaker affect, and pitch, as well as in instances of specific-language-irrelevant phonetic contrasts (e.g., Werker & Tees, 1984), there are abundant data showing that infants’ sensitivity to variations that are not functionally useful in the native language decreases across the second half of the first year of life. Such sensitivities are never completely lost; studies with adults show effects of token-specific details on memory for aurally presented words (Bradlow, Nygaard, & Pisoni, 1999; Church & Schacter, 1994; Fisher, Hunt, Chambers, & Church, 2001; Goldinger, 1998; Luce & Lyons, 1998; Nygaard, Sommers, & Pisoni, 1995; Palmeri, Goldinger & Pisoni, 1993; Pisoni, 1997). In more mature listeners, however, these effects comprise modest slowing of recognition, rather than the sort of recognition failure that has been observed in younger infants.

Availability of General Purpose Acoustic Variability

The pattern that we observed in Study 1—alternation of presence and absence of emphatic stress on repeated pronunciations of the same phonotactic sequences—ought to provide optimal evidence that emphatic stress is, in fact, not relevant for lexical identity. This is, after all, what infants (at least infants 9 months and older) hear. But our perceptual results from 7.5-month-olds challenge this interpretation; the younger infants were strongly influenced by acoustic similarity. So do younger infants hear the same thing as older infants? Extant examples of caretaker repetitions have been drawn from conversations with older infants; for example, the sequence quoted by Bernstein Ratner (1996) that was cited in the Introduction was addressed to a 13-month-old. Few studies have examined speech to very young infants, and none of these have analyzed the phonetic, prosodic, or paralinguistic properties of caretaker repetitions. But we speculate that they do; consideration of *why* caretakers repeat offers clues to *when* they repeat.

First, although caretakers’ natural demonstration of which properties do and do not covary with the phonotactic sequences that define words may be an extremely useful benefit of varying word repetition, at least in the long term, it is highly unlikely that caretakers repeat for this reason. For most lay speakers, the properties of speech that are linguistically relevant are un contemplated and self-evident. Rather than serving long-term demands of learning, repetitions serve immediate discourse needs (Fernald, 2000). Caretakers may repeat to engage or maintain an infant’s interest in an interaction. Or caretakers may repeat to draw an infant’s attention to a particular aspect of the nonlinguistic context. Caretakers may repeat to foster an infant’s comprehension. Finally, caretakers may repeat to elicit the production of a word or phrase by the infant him- or herself. Six-month-olds are not yet producing any words, so this last reason cannot be operative for them. If caretakers believe that infants are not yet capable of understanding language—probably the case with respect to the great majority of 6-month-olds—they may not be inclined to repeat or recast in order to improve infants’ comprehension.

This also applies to use of repetitions to draw infants' attention to aspects of the environment; pointing to a body part, moving an object into the infant's line of gaze, or making an object do interesting things are surely more effective means of eliciting attention than is repeating linguistic labels. Thus, some of the reasons for repeating words in varying linguistic contexts, which ensure variation in phonetic form, do not apply in interactions with very young infants. Clearly, data on speech directed to younger infants are needed to elucidate this issue.

Nonetheless, we know that some rituals with young infants do involve linguistic repetition. In our culture, the most common of these is no doubt "peek-a-boo". One hallmark of this ritual, however, is that "peek-a-boo" is repeated with a minimum of variation. Similarly, in an analysis of sentence-level repetitions, Fernald and Morikowa (1993) found a significantly higher proportion of exact repetitions in speech addressed to American and Japanese 6-month-olds than to 9- or 12-month-olds. Therefore, in the absence of additional, more direct evidence, one might conclude that the sort of repetition-with-variation that we observed in Study 1 is experienced much more often by older infants than by infants who are at the beginning stages of spoken word recognition. This is an empirical question meriting examination.

Still, why do mothers repeat in the particular pattern that we observed? Certainly there were factors in our elicitation study that may have boosted this behavior by mothers artificially (e.g., their need to engage the infant in what amounted to a relatively static and repetitive event), but we speculate that this pattern might arise from the competing pressures that mother are subjected to when conversing with their (in this case, older) infants quite often. As very well practiced speakers, mothers are used to reducing words on repeated mentions. On the other hand, if mothers wish to teach a word, or draw attention to the referent of a particular word, it is natural for them to highlight their mentions of that word. Highlighting and emphasizing are, by their very nature, contrastive: if every mention of a word is emphatic, then none is. Acceding, at least temporarily, to familiar adjustments based on the discourse status of the word provides mothers with a means of supplying the necessary contrast. Having uttered one reduced mention of the word, mothers are free to emphasize again. Of course, more nuanced analyses may reveal that mothers were shifting focus to other aspects of the event (e.g., the turtle agent) by deemphasizing the target animal's label, a view that is consistent with more recent accounts of how stress is used between adults (Dahan, Tanenhaus, & Chambers, 2002). But cycling back and forth between emphatic and nonemphatic stress might just be the simplest way to reconcile an intention to consistently emphasize a word with the habitual bias towards producing attention-getting and/or maintaining, rhythmic speech.

Having considered how, why, and when caretakers repeat, we might also consider *what* caretakers repeat. In our puppet-show procedure, before each scene, mothers were prompted with both a noun and a verb: "gazelle/push". We reported only on their repetitions of the target nouns, which were produced by mothers as either proper nouns in initial position or as count nouns in medial and final position. Given our inclusion of sentences with target nouns in all three sentence positions, it is unclear how this differentiation influences segmentation. Our data do not allow us to examine this issue. Some mothers did repeat verbs accompanying target nouns, but not nearly as often as they repeated the target nouns themselves. Analysis of repeated verbs was complicated by the fact that mothers often used them in varying morphological forms:

Look, the turtle's **pushing** the gazelle.

Ooh, the gazelle got **pushed** again.

Did the turtle **push** the gazelle?

Moreover, studies of infant word recognition have focused on recognition of nouns. An exception is a recent study by Nazzi, Dilley, Jusczyk, Shattuck-Hufnagel, & Jusczyk (2005),

who found segmentation of verbs by English-learning infants only at 13.5 months, six months later than infants are segmenting nouns (cf. Jusczyk, Houston, & Newsome, 1999). One reason for this delay in verb recognition may be that nouns are more likely to occur in sentence-initial or -final positions, positions that are privileged for segmentation and recognition (Seidl & Johnson, 2006). To date, there have been no demonstrations of infant recognition of sentence-medial words within the first year of life. For all these reasons, we believe that caretaker repetition of nouns is most significant in early word learning and language development.

Ultimately, our findings serve to inform the relation between infant preference and processing during listening. Infant preferences initially tend toward the perceptually salient, language-general (even non-linguistic) aspects of an auditory scene. These preferences include biases towards infant-directed speech (Fernald, 1985), positive affect (Mastropieri & Turkewicz, 1999; Singh et al., 2002), and higher amplitude (Sinnott, Pisoni, & Aslin, 1983), to name just a few. What is notable about each of these examples of early infant auditory preference is their acoustic similarity to emphatic stress. As we noted earlier, the results of Study 2C are consistent with a continuing preference for emphatic stress: words familiarized with emphatic stress elicited longer orientation times at recognition than words with nonemphatic stress. Nevertheless, there does not appear to be an absolute advantage for emphatic stress in early word recognition. Rather, as the results of Study 3 demonstrate, infants are forming detailed representations based on their individual encounters with words, and their recognition of new instances is constrained by the acoustic similarity among all of these details. Preference therefore does not determine processing: preferred stimulus properties enter into infants' processing, but so do many non-preferred properties.

Ultimately, to cope with variation in word forms, infants must distinguish those features that are lexically relevant from those that are not. By providing massed repetitions of words that systematically vary on lexically irrelevant dimensions, caretakers may provide input that is high optimal for learning. As we have shown, rather than abandoning the pattern of emphasis suited for more mature listeners, in speaking to their infants, mothers integrate it with the repetition that is unique to infant-directed speech. This practice may help relieve some of the stress from the formidable task that is learning to recognize spoken words.

References

- Aslin, R.; Woodward, J.; LaMendola, N.; Bever, T. Models of word segmentation in fluent maternal speech to infants. In: Morgan, J.; Demuth, K., editors. *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*. Hillsdale, NJ: Erlbaum; 1996.
- Baum S. Word recognition in individuals with left and right hemisphere damage: The role of lexical stress. *Applied Psycholinguistics* 2002;23:233–246.
- Bernstein Ratner N. Durational cues which mark clause boundaries in mother-child speech. *Journal of Phonetics* 1986;14:303–309.
- Bernstein Ratner, N. From “signal to syntax”: But what is the nature of the signal?. In: Morgan, J.; Demuth, K., editors. *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*. Mahwah, NJ: Lawrence Erlbaum; 1996.
- Bock JK, Mazzella J. Intonational marking of given and new information: Some consequences for comprehension. *Memory & Cognition* 1983;11:64–76.
- Boersma, P.; Weenink, D. Praat, a system for doing phonetics by computer. 2002. <http://www.fon.hum.uva.nl/praat/>
- Bolinger D. Accent is predictable (if you're a mind-reader). *Language* 1972;58:505–533.
- Bortfeld H, Morgan J, Golinkoff R, Rathbun K. *Mommy* and me: Familiar names help launch babies into speech stream segmentation. *Psychological Science* 2005;16:298–304. [PubMed: 15828977]
- Bradlow A, Nygaard L, Pisoni D. Effects of talker, rate, and amplitude variation on recognition memory. *Perception & Psychophysics* 1999;61:206–219. [PubMed: 10089756]

- Chafe, W. Givenness, contrastiveness, definiteness, subjects, topics, and points of view. In: Li, C., editor. *Subject and topic*. New York, NY: Academic Press; 1976.
- Church BA, Schacter DL. Perceptual specificity of auditory priming: Implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1994;20:521–533.
- Clark, HH.; Haviland, SE. Comprehension and the given-new contract. In: Freedle, R., editor. *Discourse production and comprehension*. Norwood, NJ: Ablex; 1977.
- Cooper RP, Aslin R. Preference for infant-directed speech in the first month after birth. *Child Development* 1990;61:1584–1995. [PubMed: 2245748]
- Cutler, A. *Cognition models of speech processing: Psycholinguistic and computational perspectives*. Cambridge, MA: MIT Press; 1990. Exploiting prosodic probabilities in speech segmentation.
- Cutler A, Butterfield S. Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language* 1992;31:218–236.
- Cutler A, Norris D. The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance* 1988;14:113–121.
- Cutler A, Swinney D. Prosody and the development of comprehension. *Journal of Child Language* 1987;14:145–167. [PubMed: 3558521]
- Cutler A, Dahan D, van Donselaar W. Prosody in the comprehension of spoken language: A literature review. *Language and Speech* 1997;40:141–201. [PubMed: 9509577]
- Dahan D, Tanenhaus M, Chambers C. Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language* 2002;47:292–314.
- D’Odorico L, Jacob V. Prosodic and lexical aspects of maternal linguistic input to late-talking toddlers. *International Journal of Language and Communication Disorders* 2006;41:293–311. [PubMed: 16702095]
- Dunlop WP, Cortina JM, Vaslow JB, Burke MJ. Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods* 1996;1:170–177.
- Echols C, Crowhurst M, Childers J. The perception of rhythmic units in speech by infants and adults. *Journal of Memory and Language* 1997;36:202–225.
- Ferguson CA. Baby talk in six languages. *American Anthropologist* 1964;66:103–114.
- Fernald, A. The perceptual and affective salience of mothers’ speech to infants. In: Feagans, L.; Garvey, C.; Golinkoff, R., editors. *The origins and growth of communication*. Norwood, NJ: Ablex; 1984. p. 5-29.
- Fernald A. Four month old infants prefer to listen to motherese. *Infant Behavior and Development* 1985;8:181–195.
- Fernald A. Speech to infants as hyper-speech: Knowledge-driven processes in early word recognition. *Phonetica* 2000;57:242–254. [PubMed: 10992144]
- Fernald A, Kuhl P. Acoustic determinants of infant perception for motherese speech. *Infant Behavior and Development* 1987;10:279–293.
- Fernald A, Mazzie C. Prosody and focus in speech to infants and adults. *Developmental Psychology* 1991;27:209–221.
- Fernald A, Morikawa H. Common themes and cultural variations in Japanese and American mothers’ speech to infants. *Child Development* 1993;64:637–656. [PubMed: 8339686]
- Fernald A, Simon T. Expanded intonation contours in mothers’ speech to newborns. *Developmental Psychology* 1984;20:104–113.
- Fernald A, Taeschner T, Dunn J, Papousek M, de Boysson-Bardies B, Fukui I. A cross-language study of prosodic modification in mothers’ and fathers’ speech to preverbal infants. *Journal of Child Language* 1989;16:477–501. [PubMed: 2808569]
- Fisher C, Hunt C, Chambers K, Church B. Abstraction and specificity in preschoolers’ representations of novel spoken words. *Journal of Memory and Language* 2001;45:665–687.
- Fisher C, Tokura H. The given/new contract in speech to infants. *Journal of Memory and Language* 1995;34:287–310.
- Fowler C. Differential shortening of repeated content word produced in various communicative contexts. *Language and Speech* 1988;28:47–56.

- Fowler C, Housum J. Talkers' signaling of 'new' and 'old' words in speech and listeners' perception and use of that distinction. *Journal of Memory and Language* 1987;26:489–504.
- Fowler C, Levy E, Brown J. Reductions of spoken words in certain discourse contexts. *Journal of Memory and Language* 1997;37:24–20.
- Goldinger S. Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1998;22:1166–1183.
- Goodman, J.; Nusbaum, H.; Lee, L.; Broihier, K. The effects of syntactic and discourse variables on the segmental intelligibility of speech. *Proceedings of the International Conference on Spoken Language Processing.*; 1990. p. 393-396.
- Gravano, A.; Hirschberg, J. Effect of genre, speaker, and word class on the realization of given and new information. *Proceedings of Interspeech 2006*; Pittsburgh, PA. September; 2006.
- Halliday MAK. Notes on transitivity and theme in English: II. *Journal of Linguistics* 1967;3:199–244.
- Haviland S, Clark H. What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior* 1974;13(5):512–521.
- Höhle, B.; Weissenborn, J. Discovering grammar: Prosodic and morpho-syntactic aspects of rule formation in first language acquisition. In: Friederici, A.; Randolph, M., editors. *Learning: Rule extraction and representation*. Berlin, Germany: Walter de Gruyter & Co; 1999. p. 37-69.
- Houston D, Jusczyk P. The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance* 2000;26:1570–1582. [PubMed: 11039485]
- Houston D, Jusczyk P, Kuijpers C, Coolen R, Cutler A. Cross-language word segmentation by 9-month-olds. *Psychonomic Bulletin & Review* 2000;7:504–509. [PubMed: 11082857]
- Johnson E, Jusczyk P. Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language* 2001;44:548–567.
- Jusczyk P, Aslin R. Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology* 1995;29:1–23. [PubMed: 7641524]
- Jusczyk P, Houston D, Newsome M. The beginning of word segmentation in English-learning infants. *Cognitive Psychology* 1999;39:159–207. [PubMed: 10631011]
- Kemler Nelson D, Jusczyk P, Mandel D, Myers J, Turk A, Gerken L. The headturn preference procedure for testing auditory perception. *Infant Behavior and Development* 1995;18:111–116.
- Ladd, DR. *The structure of intonational meaning*. Bloomington, IN: Indiana University Press; 1980.
- Ladd, DR. *Intonational phonology*. Cambridge, England: Cambridge University Press; 1996.
- Luce P, Lyons E. Specificity of memory representations for spoken words. *Memory and Cognition* 1998;26:708–715.
- Mandel-Emer, D. *Dissertation Abstracts International*. Vol. 57. 1997. Names as early lexical candidates: Helpful in language processing? (Doctoral dissertation, State University of New York, Buffalo, 1997); p. 5947
- Mastropieri D, Turkewicz G. Prenatal experience to neonatal responsiveness to vocal expressions of emotion. *Developmental Psychobiology* 1999;35:204–214. [PubMed: 10531533]
- Mattys S, Jusczyk P, Luce P, Morgan J. Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology* 1999;38:465–494. [PubMed: 10334878]
- Mattys S, Samuel A. How lexical stress affects speech segmentation and interactivity: Evidence from the migration paradigm. *Journal of Memory and Language* 1997;36:87–116.
- Morgan J. A rhythmic bias in preverbal speech segmentation. *Journal of Memory and Language* 1996;35:666–688.
- Morgan J, Saffran J. Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development* 1995;66:911–936. [PubMed: 7671658]
- Nazzi T, Dilley L, Jusczyk AM, Shattuck-Hufnagel S, Jusczyk P. English-learning infants' segmentation of verbs from fluent speech. *Language and Speech* 2005;48:279–298. [PubMed: 16416938]
- Nazzi T, Iakimova G, Bertoni J, Frédonie S, Alcantara C. Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. *Journal of Memory and Language* 2006;54:283–299.

- Nooteboom SG, Kruyt JG. Accents, focus, distribution, and the perceived distribution of given and new information: An experiment. *Journal of the Acoustical Society of America* 1987;82:1512–1524. [PubMed: 3693693]
- Norris D, McQueen J, Cutler A. Competition and segmentation in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory & Cognition* 1995;21:1209–1228.
- Nygaard L, Sommers S, Pisoni D. Speech perception as a talker-contingent process. *Psychological Science* 1994;5:42–46.
- Palmeri T, Goldinger S, Pisoni D. Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1993;19:309–328.
- Pisoni, D. Some thoughts on “normalization” in speech perception. In: Johnson, K.; Mullennix, JW., editors. *Talker variability in speech processing*. San Diego, CA: Academic Press; 1997.
- Prince, EF. Toward a taxonomy of given-new information. In: Cole, P., editor. *Radical pragmatics*. New York, NY: Academic Press; 1981. p. 223-255.
- Rochemont, M.; Culicover, P. *English focus constructions and the theory of grammar*. New York, NY: Cambridge University Press; 1990.
- Saffran J, Aslin R, Newport E. Statistical learning by 8-month-old infants. *Science* 1996;274:1926–1928. [PubMed: 8943209]
- Seidl A, Johnson E. Infant word segmentation revisited: Edge alignment facilitates target extraction. *Developmental Science* 2006;9:565–573. [PubMed: 17059453]
- Selkirk, EO. *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press; 1984.
- Shattuck-Hufnagel S, Turk A. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 1996;25:193–247. [PubMed: 8667297]
- Singh L. Influences of high and low variability on infant word recognition. *Cognition* 2008;106:833–870. [PubMed: 17586482]
- Singh L, Morgan J, Best C. Infants’ listening preferences: Baby talk or happy talk? *Infancy* 2002;3:365–394.
- 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;51:173–189.
- Singh L, Nestor S, Bortfeld H. Overcoming the effects of variation in infant speech segmentation: Influences of word familiarity. *Infancy* 2008;13:57–74.
- Singh L, White K, Morgan J. Building a lexicon in the face of variable input: Effects of pitch and amplitude variation on early word recognition. *Language Learning and Development* 2008;4:157–178.
- Sinnott J, Pisoni D, Aslin R. A comparison of pure tone auditory thresholds in human infants and adults. *Infant Behavior and Development* 1983;6:3–17.
- Slowiazcek L. Effects of lexical stress in auditory word recognition. *Language and Speech* 1990;33:47–68. [PubMed: 2283920]
- Small L, Simon S, Goldberg J. Lexical stress and lexical access: Homographs versus nonhomographs. *Perception & Psychophysics* 1988;44:272–280. [PubMed: 3174358]
- Snow CE. Mothers’ speech to children learning language. *Child Development* 1972;43:549–565.
- Solan L. Contrastive stress and children’s interpretation of pronouns. *Journal of Speech & Hearing Research* 1980;23:688–698. [PubMed: 7421166]
- Terken J, Nooteboom SG. Opposite effects of accentuation and deaccentuation on verification latencies for Given and New information. *Language and Cognitive Processes* 1987;2:145–163.
- Thiessen E, Hill Saffran J. Infant-directed speech facilitates word segmentation. *Infancy* 2005;7:53–71.
- Vroomen J, Tuomainen J, de Gelder B. The role of word stress and vowel harmony in speech segmentation. *Journal of Memory and Language* 1998;38(2):133–149.
- Werker J, McLeod P. Infant preference for both male and female infant-directed-talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology* 1989;43:230–246. [PubMed: 2486497]
- Werker J, Tees RC. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 1984;7:49–63.

Appendix

Appendix A. Development of Emphatic and Nonemphatic Stimuli

Here we explain the method used to develop emphatic and nonemphatic utterances. The stimuli were modeled on the spontaneous utterances produced by mothers in Study 1. From this corpus of infant-directed speech, we chose 120 utterances. Of these, forty had a target content word (i.e., a strong-weak animal's name) in initial position, 40 had a target content word in medial position, and 40 had a target content word in final position. These utterances were sampled into sound files that were then concatenated in random order. Twenty English-speaking adults (9 males and 11 females) listened to the 120 digitized sentences extracted from the transcripts. Participants first received a brief tutorial about the difference between emphatic and nonemphatic stress, and heard several examples of each in utterances spliced from naturally produced fluent speech.

Using a 7-point Likert scale, listeners rated each of the 120 sentences for how emphatic the target word embedded in it sounded (with 1 anchored at "very unemphatic" and 7 anchored at "very emphatic"). The target word within each utterance was underlined in the transcription, indicating to participants which word was to be rated. Participants first rated several practice sentences before rating the 120 test sentences. Each sentence was played twice, with a two-second pause between repetitions and participants were allowed to take as much time as they wanted to rate the word before the next sentence was played.

Collapsing across target word stress type, ratings for sentence position were: 2.15 (SD = 1.18) for initial position, 3.16 (SD = .89) for medial position, and 3.43 (SD = 1.52) for final position (where neutral stress was represented by the intermediate point – 4 – on the scale). The slightly lower ratings for target words in initial position reflect the tendency for words in this position – usually sentential subjects – to be acoustically reduced relative to targets in other positions. This highlights the importance of controlling for sentence position when manipulating emphatic stress.

Using these ratings, we identified equal numbers of sentences that had targets in each of the three sentence positions and that were also rated within 2 points of the scale endpoints. This resulted in 72 sentences, with half containing target words produced with emphatic stress and half containing target words produced with nonemphatic stress. Within each set of 36, there were equal numbers of targets in each of the three possible sentence positions. The mean rating for selected emphatic stress sentences was 6.28 (SD = .97) and for selected nonemphatic stress sentences was 1.57 (SD = .89). These sentences served as templates for constructing single-speaker stimuli for use in the perceptual experiments. Individual words were selected from the collection of sentences, isolated, and submitted to the same rating process as that described above. The 30 instances of each that received the highest and lowest scores were selected as nonemphatic and emphatic exemplars, respectively. The mean rating for selected emphatically stressed words was 6.36 (SD = .86) and for selected nonemphatically stressed words was 1.71 (SD = .91).

To eliminate talker differences across the stimuli, we recorded a native English-speaking female, who mimicked the original intonation pattern for each sentence template while replacing the animal name in the original sentence with each of the four animal names selected for use in the perceptual studies. The speaker also mimicked the word exemplars for each of the four animal names. The speaker recorded the words and sentences while addressing her own infant, who was with her in the recording booth. Each item was played to and repeated by our speaker twice. She was instructed to follow the intonation pattern of the original version as closely as possible. The version judged by the first author (HB) to sound more like the original was retained for use in the perceptual studies.

Several measures were undertaken to ensure that these mimicked stimuli reflected the form of stress in the original items. As before, the sentences were judged for type of emphatic stress (e.g., whether target words were stressed emphatically or nonemphatically) by English-speaking adults who were naïve to the hypothesis being tested. These participants received a brief tutorial on the difference between emphatic and nonemphatic stress and completed several practice trials before beginning. They then heard each sentence two times. As in the earlier rating session, there was a two second pause between repetitions and each participant was allowed to take as long as they wanted before hearing the next sentence. Responses for stress type were considered correct if the answer matched the form of stress intended during recording. The average consistency between these ratings and the type of stress indicated by the Likert scale ratings collected earlier was 91.23%. Items eliciting unreliable responses were re-recorded by the same female speaker and re-rated until all were judged reliably. Sentences were then digitally arranged to create test stimuli for the perceptual experiments reported here.

Acoustic measures for the final set of words used as familiarization stimuli are shown in Table A1. Analyses of these stimuli were consistent with the acoustic measures of the naturally produced, infant-directed speech reported in Study 1. First, maximum F_0 was higher for emphatically stressed words than for nonemphatically stressed words $t(59) = 22.20, p < .0001$, but minimum F_0 was not, $t(59) = 1.57, ns$. Overall, mean F_0 was higher in emphatically stressed words than in nonemphatically stressed words, $t(59) = 18.01, p < .0001$. Finally, pitch range (in semitones) of emphatic stressed words exceeded that of nonemphatically stressed words, $t(59) = 16.12, p < .0001$, as did relative durations of targets produced with the two forms of stress $t(59) = 19.28, p < .0001$.

Appendix B. Example Sentence Sets: Alternating (emphatic/nonemphatic) stress across mentions

He's nudging the <u>chicken</u>	<u>Falcon</u> is saying 'hello'!
Oh I think <u>chicken</u> is tough to push	He's rocking the <u>falcon</u>
He's lifting the <u>chicken</u> right up in the air	<u>Falcon</u> is getting swung all around
<u>Chicken</u> is being rocked	There's that <u>falcon</u> again
He's tickling the <u>chicken</u>	He's swinging that <u>falcon</u> round and round
<u>Chicken</u> is all gone	Look at the <u>falcon</u>
Now he's pulling the <u>monkey</u>	<u>Dolphin</u> is rolling over
<u>Monkey</u> looks heavy	Look at that <u>dolphin</u> rolling and rolling
I don't think you've seen a <u>monkey</u>	He's pulling the <u>dolphin</u>
<u>Monkey</u> is being tickled	Rock-a-bye <u>dolphin</u>
We have a <u>monkey</u> on our puzzle	<u>Dolphin</u> looks tired
He pushed the <u>monkey</u> right off	Oh, the <u>dolphin</u> is saying 'bye bye'!

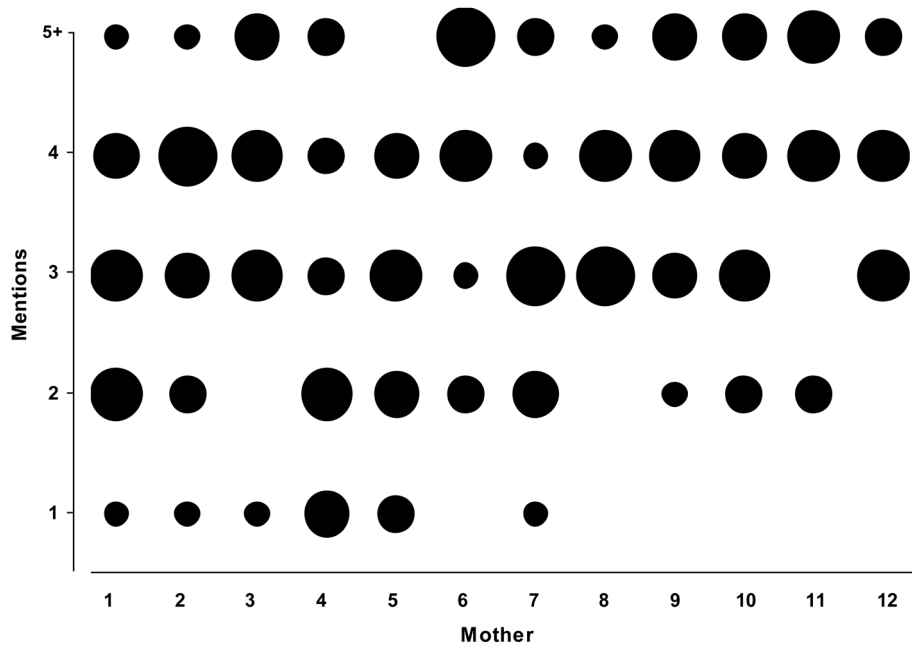


Figure 1. Study 1: Graphic presentation of number of mentions by each mother across target words. The area of each dot is proportional to the number of words that a mother mentioned a given number of times. Mother 7 mentioned one word once, three words twice, five words three times, one word four times, and two words five or more times.

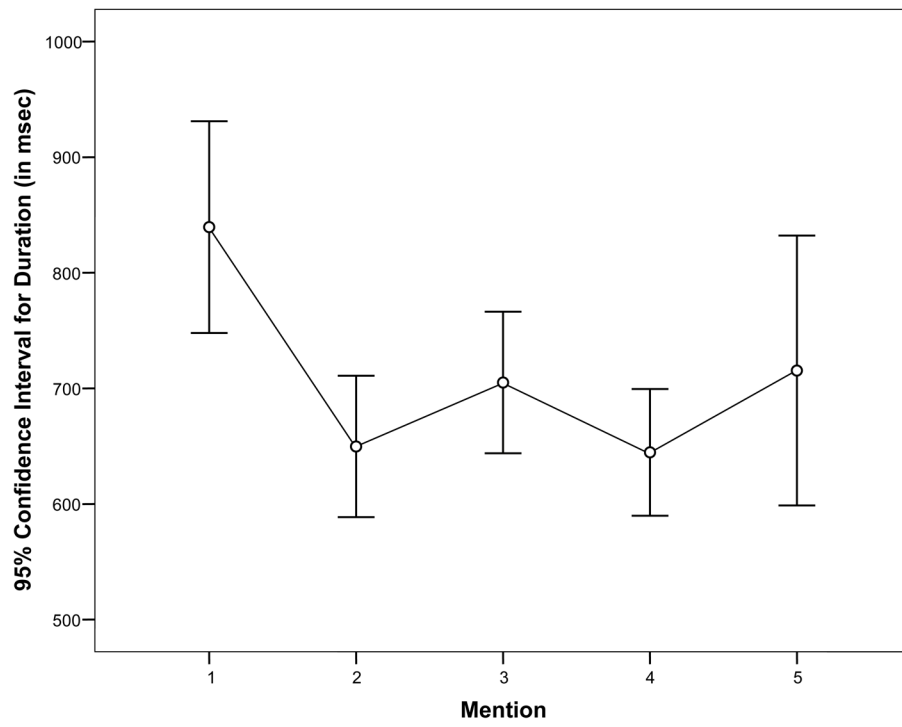


Figure 2.
Study 1: Average duration of target words across five mentions (95% confidence intervals).

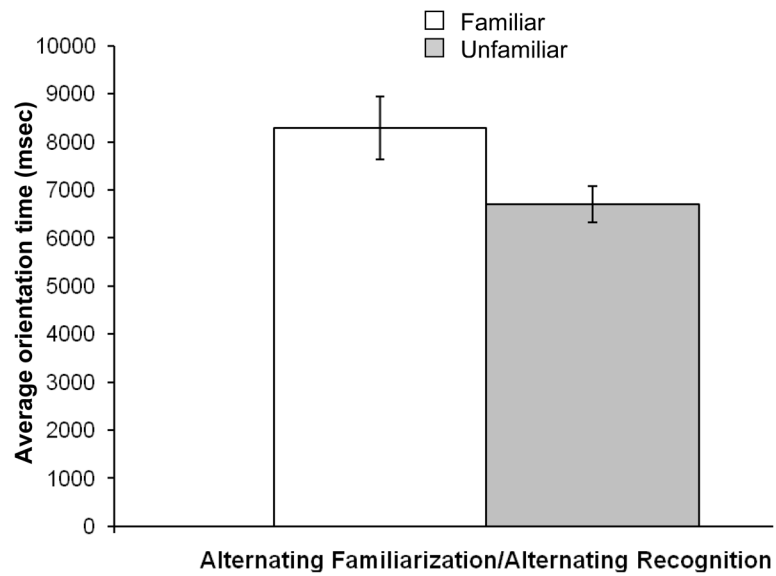


Figure 3. Study 2A: Recognition orientation times (within-subjects): alternating (emphatic/nonemphatic) familiarization stress with alternating (emphatic/nonemphatic) stress in recognition sentence sets.

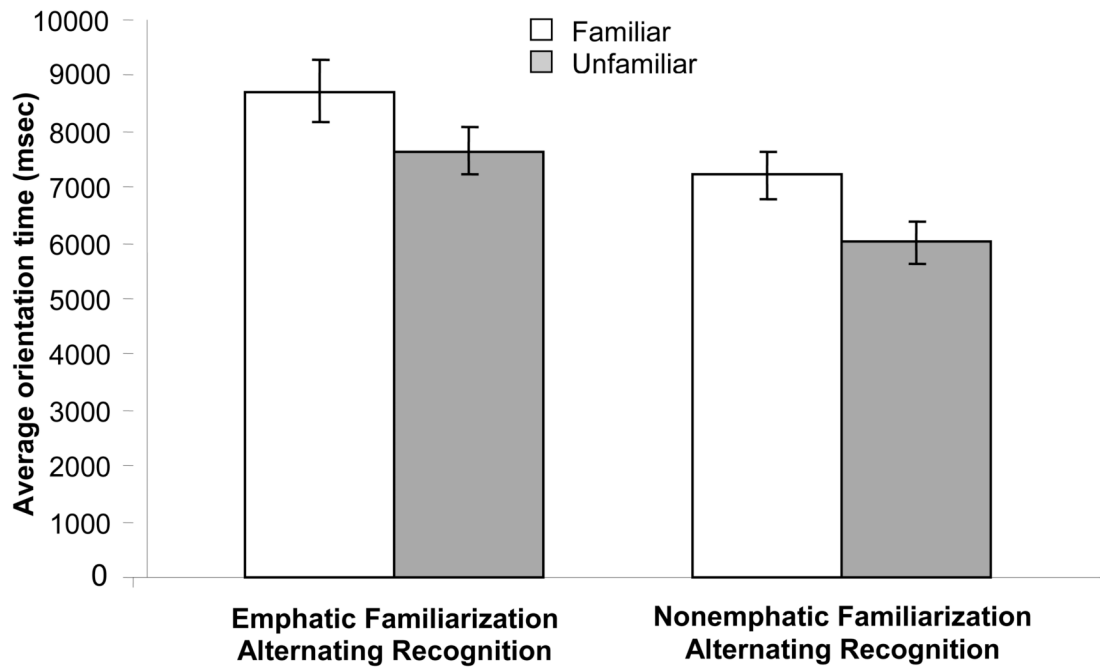


Figure 4. Study 2B: Recognition orientation times (between-subjects): 1) emphatic familiarization stress and alternating (emphatic/nonemphatic) stress in recognition sentences OR 2) nonemphatic familiarization stress with alternating (emphatic/nonemphatic) stress in recognition sentence sets.

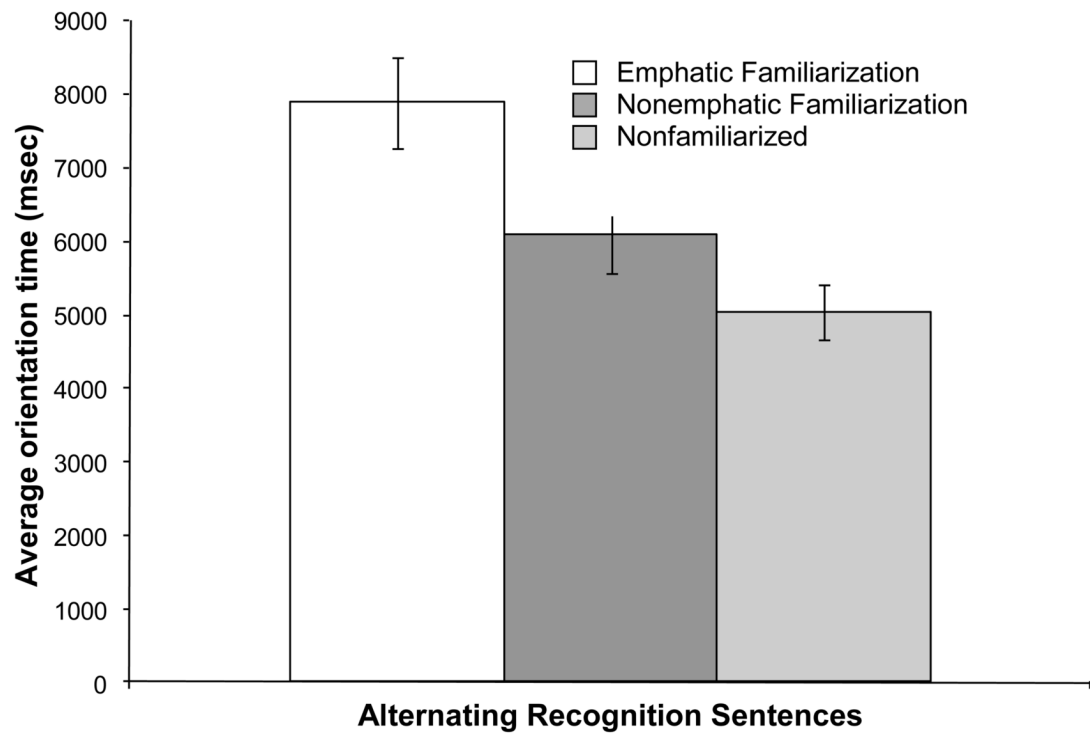


Figure 5. Study 2C: Recognition orientation times (within-subjects): 1) emphatic familiarization stress on one word and 2) nonemphatic familiarization stress on another word with alternating stress in recognition sentence sets (within-subjects).

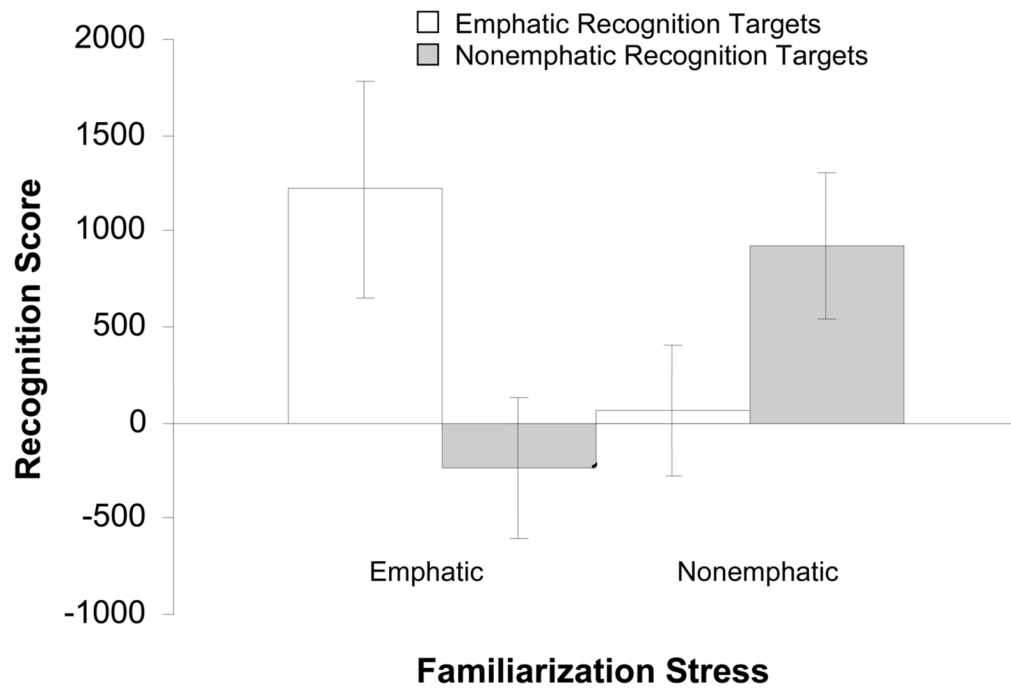


Figure 6. Study 3: Recognition scores (between-subjects): Recognition orientation time differences for sentence sets containing familiarized words minus sentence sets containing nonfamiliarized words for matched vs. unmatched stress from familiarization to recognition.

Table 1

Bisyllabic Animal Names from Puppet Show

Initial Stress	Non-Word-Initial Stress
Monkey	Giraffe
Walrus	Baboon
Chicken	Gazelle
Zebra	Raccoon

Table 2
Acoustic Analyses of Naturally Produced Words: Mean and (Standard Deviations)

By Mention	Mean F ₀	Min F ₀	Max F ₀	F ₀ Range	Duration
First (N = 185)	305.75 (79.36)	200.64 (67.13)	410.85 (113.30)	12.59 (5.14)	842.22 (293.90)
Second (N = 177)	287.65 (84.18)	216.90 (74.97)	358.40 (115.06)	8.66 (5.25)	652.10 (234.91)
Third (N = 151)	306.96 (86.23)	213.55 (72.93)	400.37 (121.16)	11.01 (6.42)	711 (279.27)
Fourth (N = 97)	287.14 (82.24)	214.94 (76.05)	359.35 (110.76)	9.11 (5.80)	636.19 (193.74)
Fifth (N = 38)	319.01 (92.84)	214.04 (76.20)	423.97 (144.17)	11.75 (7.40)	686.99 (223.44)
Overall (N = 648)	299.08 (83.35)	211.02 (72.65)	387.14 (119.53)	10.58 (5.94)	719.94 (269.72)

Note: Mean, minimum, and maximum F₀ in Hertz, F₀ range in semitones, and duration in milliseconds.

Table 3

Means and Standard Deviations for Interaction in Study 3

Stress at Familiarization	Between-Subjects Variables		Within-Subjects Variable				
	<i>Stress at Recognition</i>	<i>Familiar</i>	<i>Familiar</i>	<i>Unfamiliar</i>	Difference	Significance	Effect Size
	Mean	SD	Mean	SD			
Emphatic	7391	2881	6174	1609	1217	0.05	0.19
Nonemphatic	6670	1777	6912	1914	-243	0.52	0.02
Emphatic	7290	1284	7226	1815	63	0.86	0.00
Nonemphatic	6371	1812	5445	2051	927	0.03	0.24

Note: Mean looking time (measured in milliseconds) based on stress at recognition and familiarization (between-subject variable) and word familiarity (within-subject variable). Significant differences in looking time due to familiarization indicate word recognition. Effect sizes are computed as partial η^2 .

Table A1

Acoustic Analyses of Words: Means and (Standard Deviations)

	Mean F ₀	Min F ₀	Max F ₀	F ₀ Range	Duration
Emphatic	272.95 (45.07)	151.81 (28.01)	434.44 (82.38)	16.14 (4.02)	782.52 (115.94)
Nonemphatic	171.33 (24.23)	144.84 (28.77)	198.57 (25.87)	5.62 (2.68)	565.86 (81.36)

Note: Mean, minimum, and maximum F₀ in Hertz, F₀ range in semitones, and duration in milliseconds.