INFANT BEHAVIOR AND DEVELOPMENT **19,** 401–418 (1996)

Stressed and Word-Final Syllables in Infant-Directed Speech

DREMA DIAL ALBIN AND CATHARINE H. ECHOLS
University of Texas at Austin

Comparisons between infant-directed and adult-directed speech were conducted to determine whether word-final syllables are highlighted in infant-directed speech. Samples of adult-directed and infant-directed speech were collected from 8 mothers of 6-month-old and 8 mothers of 9-month-old infants. Mothers were asked to label seven objects both to an experimenter and to their infant. Duration, pitch, and amplitude were measured for whole words and for each of the target word syllables. As in prior research, the infant-directed targets were higher pitched and longer than adult-directed targets. The results also extend beyond previous results in showing that lengthening of final syllables in infant-directed speech is particularly exaggerated. Results of analyses comparing word-final versus nonfinal unstressed syllables in utterance-medial position in infant-directed speech showed that lengthening of unstressed word-final syllables occurs even in utterance-internal positions. These results could suggest a mechanism for proposals that word-final syllables are perceptually salient to young children.

infant-directed speech motherese word-level segmentation language acquisition

Increasingly, researchers in the field of language acquisition have come to believe that prelinguistic biases may assist language learning by enabling infants to identify linguistically relevant elements of speech (Gleitman, Gleitman, Landau, & Wanner, 1988; Gleitman & Wanner, 1982; Peters, 1983; Slobin, 1973, 1985). One set of biases may assist with the identification of words in speech. Although there is no single consistent set of acoustic markers for word boundaries in English (Cole & Jakamik, 1980; Hayes & Clark, 1970), children may be assisted by a tendency to attend to and extract syllables that are highlighted by stress, pitch, or other aspects of prosody (Gleitman & Wanner, 1982). Such a tendency would reduce the difficult problem of identifying boundaries between words to the simpler task of identifying highlighted syllables. However, this tendency would be valuable only if the input provides cues that are in fact advantageous for identifying words. Accordingly, it is important to describe the patterns of prosodic highlighting in infant-directed speech and, in particular, to determine whether the highlighting provides useful cues to word-level units.

Before turning to the question of whether particular syllables are prosodically highlighted in infant-directed speech, we will discuss stress in adult-directed speech, characteristics of infant-directed speech, and young children's sensitivity to prosodic cues in speech. A rather extensive body of research describes prosodic differences between infant-directed and adult-directed speech (see Fernald, 1991, for a review). There is also a rapidly expanding literature describing infants' sensitivity to prosodic cues in their native language (Hirsh-Pasek et al., 1987; Jusczyk et al., 1992; Mehler et al., 1988).

STRESS IN ADULT-DIRECTED SPEECH

An acoustic definition of stress is notoriously elusive (Hayes, 1985; Lehiste, 1970). In adult-directed English, stressed syllables are distinguished by one or more of the following: higher amplitude, or loudness; higher fundamental frequency, or pitch; nonreduced vowel quality; and greater duration (Lehiste, 1970). Any of these

We would like to thank Randy Diehl and Janet Spence for helpful critical readings of earlier versions of this article and three anonymous reviewers for valuable suggestions. Gratitude is also extended to Hollie Davis for assistance in data collection; to Marlena Creusere, Catherine Malerba, L'Don Sawyer, and Theresa Soliz for assistance with acoustical analyses; to Jerry Manheimer for statistical advice; and to the mothers and infants who participated. Portions of this research were presented at the International Conference on Infant Studies, Paris, France, June, 1994. This research was supported by Grant #003658-368 from the Advanced Research Program of the Texas Higher Education Coordinating Board and by a research grant from the University Research Institute at the University of Texas to C.H.E.

Correspondence and requests for reprints should be sent to Catharine Echols, Department of Psychology, 330 Mezes Hall, University of Texas, Austin, TX 78712.

features may contribute to the perception of stress, although some features (particularly pitch and duration) appear to be more dominant.

Stress tends to be increased for focused words within a sentence. For example, when introduced as new information, words may be highlighted by increased pitch (Brown, 1983), and they are particularly likely to occur on pitch peaks (Fernald & Mazzie, 1991). The highlighting provided by stress may influence speech processing (Grosjean & Gee, 1987), contributing to faster recognition of elements in stressed than in unstressed syllables (Cutler & Foss, 1977). More generally, stress appears to play a role in the identification of word-level units by adults (Cutler & Butterfield, 1992; Cutler & Norris, 1988; Nakatani & Schaffer, 1978).

CHARACTERISTICS OF INFANT-DIRECTED SPEECH

As a substantial body of research now documents, infant-directed speech (also called *motherese* or *parentese*) typically differs from adult-directed speech on a variety of dimensions (see Fernald, 1991, for a review). Utterances directed to young children generally include most or all of the following properties: higher pitch, greater pitch range, shorter total length, slower rate of speech, and longer pauses (Fernald & Mazzie, 1991; Fernald & Simon, 1984; Fernald et al., 1989; Garnica, 1977; Grieser & Kuhl, 1988). Some properties of infant-directed speech may be valuable for word identification. Pauses occur more reliably at sentence boundaries (Fernald & Simon, 1984) and somewhat more reliably at phrase boundaries (Morgan, 1986) than in adult-directed speech. Analyses of monosyllables in speech directed to preverbal children show lengthening of vowels in words preceding a clause boundary (Bernstein Ratner, 1986) and exaggeration of phrase-final lengthening (Morgan, 1986). Although clause and phrase boundaries alone are insufficient for the segmentation of word-level units from speech, they may aid by providing a boundary either before or after a word.

EVIDENCE FOR INFANTS' AND YOUNG CHILDREN'S SENSITIVITY TO PROSODY

The special qualities of infant-directed speech will be useful only to the extent that a learner is prepared to pick up on the linguistically relevant properties (Gleitman et al., 1988). There is evidence that infants may come into the world attentive to prosodic characteristics of the ambient language, such as the rhythm or pitch pattern. Newborns as young as 4 days have been shown to distinguish the prosody of their native language from that of a nonnative language (Mehler et al., 1988). Infants also are sensitive to stress pattern very early: By 1 to 4 months of age, infants can discriminate changes in stress pattern (Jusczyk & Thompson, 1978; Spring & Dale, 1977). By 9 months, English-hearing infants prefer disyllabic words with a stress pattern common in English (a *trochaic*, or strong–weak, pattern) over those with a less characteristic pattern (i.e., *iambic*, or weak–strong; Jusczyk, Cutler, & Redanz, 1993). More generally, by 9 months or earlier, infants from English-speaking homes are sensitive to prosodic cues to clause and phrase boundaries in English (Hirsh-Pasek et al., 1987; Jusczyk et al., 1992; Kemler Nelson, Hirsh-Pasek, Jusczyk, & Wright Cassidy, 1989). Infant-directed speech appears to be particularly rich in these prosodic cues, as is suggested by Kemler Nelson et al.'s (1989) finding that 7- and 10-month-old infants were sensitive to prosodic cues to clause boundaries in infant- but not in adult-directed speech.

Infants' first speech productions also may provide a glimpse into which properties of the input are salient to the youngest language learners. One striking characteristic of early productions is that stressed syllables of the target word, as well as unstressed word-final syllables, are preferentially included: Children show a strong tendency to omit from their early utterances unstressed syllables that are in a nonfinal position in the target word (Allen & Hawkins, 1980; Echols & Newport, 1992; Ingram, 1978; Klein, 1981, 1984; Pollock, Brammer, & Hageman, 1989). Similar omissions are observed in children's imitations of nonsense-syllable strings and of isolated nonsense words (Blasdell & Jensen, 1970; Hura & Echols, 1996).

One interpretation of these findings is that infants are innately biased to attend to and extract stressed syllables from the speech stream. A perceptual bias of this nature could facilitate word-level segmentation by providing infants with a starting point for word identification: Infants could extract stressed syllables as their initial representations for words (Gleitman & Wanner, 1982). The tendency of word-final,

and not only stressed, syllables to be preserved in children's earliest productions is a puzzle for the perceptual biases account. Because stressed syllables typically are louder, longer, and higher pitched, there is an acoustic basis for the perceptual salience of those syllables. An acoustic basis for the perceptual salience of final syllables is less evident. It is possible, however, that an evaluation of the acoustic characteristics of stress in infant-directed speech could provide evidence of a prosodic basis for the salience of final syllables. In particular, it is conceivable that final syllables will carry some characteristics of stress as a result of the exaggerated prosody of infant-directed speech.

Characteristics of stress and of final syllables in infant-directed speech are pertinent to another account for the form of children's early productions. An alternative to the perceptual salience account described here is an account in which children omit unstressed nonfinal syllables because those syllables fail to adhere to a trochaic (strong–weak) template (Allen & Hawkins, 1980; Gerken, 1990, 1994a, 1994b see also Demuth, 1996, for a related account). On this metrical account, the advantage for final syllables can be explained because final unstressed syllables, unlike nonfinal unstressed syllables, typically will be part of a trochaic unit. Consistent with this view, English-learning children are more likely to produce unstressed syllables that conform to a trochaic sequence than unstressed syllables that are not part of a trochee (Allen & Hawkins, 1980; Gerken, 1990, 1994a, 1994b; Hura & Echols, 1996; see also Demuth, 1992; Fikkert, 1994; Wijnen, Krikhaar, & den Os, 1994, for similar evidence in Sesotho and Dutch).

The "trochaic bias" was originally proposed as a production-based phenomenon, that is, a preference for producing trochaic over iambic sequences (Allen & Hawkins, 1980; Gerken, 1990), perhaps due to a trochaic production template or to metrical considerations (Gerken, 1994a, 1994b). However, there also could be a perceptual component to the trochaic bias (Echols, 1996; Hura & Echols, 1996; Jusczyk et al., 1993): If children expect words to conform to a trochaic sequence, then they would notice and extract unstressed syllables following a stressed syllable. In contrast, they would ignore unstressed syllables prior to a stressed syllable, expecting that those syllables are not part of the word being extracted. Supporting this possibility are findings that infants may perceive trochaic sequences as more coherent than iambic sequences (Echols, Crowhurst, & Childers, in press; Morgan, 1996) and that they are more likely to extract whole trochaic sequences than whole iambic sequences from longer strings of speech (Echols et al., in press; Newsome & Jusczyk, 1995).

Regardless of whether the trochaic stress pattern is used for segmentation or whether it serves solely as a template for productions, children will have to identify the rhythmic pattern. In particular, if a trochaic segmentation strategy is to aid in the initial identification of words, then it should be available to children quite early. The previously mentioned preference for trochaic over iambic words appears to develop between 6 and 9 months (Jusczyk et al., 1993). This finding may suggest that 9-month-old infants already have identified the trochaic sequence as a typical stress pattern of the native language. If, as Jusczyk et al. (1993) suggested, the trochaic preference develops as a result of experience with the native language, then their findings raise the question of how infants could identify this stress pattern as typical of English before they comprehend more than a few words. The results of the present study also have the potential to contribute to an explanation for this puzzle in that there may be characteristics of infant-directed speech that facilitate the identification of trochees.

An investigation of characteristics of stressed and of final syllables, then, may provide information pertaining to two interrelated questions in early word learning. The first of these questions is that of why word-final syllables tend to be preserved in early productions even when they are unstressed. If word-final syllables tend to be highlighted in infant-directed speech, then that finding would assist the perceptual salience account by providing an acoustic basis for the advantage of final over nonfinal unstressed syllables. Of course, such a finding would not show that the perceptual salience account is correct; it would only contribute to a plausible explanation for the retention of final syllables in early productions. This investigation could also be informative for a second question, that of how English-hearing infants can identify the trochaic stress pattern as characteristic of words in the ambient language. The highlighting of final syl-

lables in infant-directed speech could assist children in identifying the trochaic stress pattern: When combined with the acoustic prominence of stressed syllables and the penultimate stress typical of English, such highlighting could direct attention to trochees. Alternatively, if any such highlighting meant that final syllables were perceived as stressed syllables, then it could disrupt the identification of the trochaic pattern.

CHARACTERISTICS OF STRESSED AND FINAL SYLLABLES IN INFANT-DIRECTED SPEECH

Stress appears to be a salient feature of speech both for adults and for children. Although prosodic features of infant-directed speech (including pitch levels, pitch change, durational features, and pauses) have been described, stress in infant-directed speech has been investigated less frequently. Moreover, little attention has been paid to final syllables in infant-directed speech, despite the fact that word-final syllables, like stressed syllables, are preferentially included in early productions.

The question of central interest in this study was whether characteristics of stress are expressed differently in infant- versus adult-directed speech, and, in particular, whether final syllables, including unstressed final syllables, carry some characteristics of stress. Sixteen mothers were recorded with their 6- or 9-month-old infants during an hour-long session at the participant's home. To provide some consistency across participants and across infant-directed and adult-directed contexts, mothers participated in a labeling game in which they were given seven objects with names of varying length. Mothers were asked to introduce these objects both to the experimenter and to the infant. Using this approach, labels of different numbers of syllables and with different stress patterns could be elicited in a naturalistic context. Although all of the target words elicited in this way were nouns, the tendency of object names to predominate in early speech (Gentner, 1982; Nelson, 1973) suggests that the identification of nouns is a major focus of early word identification. Adult-directed and infant-directed variants of the target words were then compared acoustically. We hypothesized that final syllables, including those that are underlyingly unstressed, may take on certain characteristics of stress in infant-directed speech. Based on prior research, we also expected that words would be higher in pitch and longer in duration when directed to infants.

METHOD

Participants

The participants were 16 mothers, with an age range of 26 to 39 years. The mothers selected for observation were the primary caregivers. Eight mothers had 6-month-old infants, and 8 had 9-month-old infants. Names of possible participants were obtained through birth announcements in a local newspaper, according to the date of the infant's birth. The parents were contacted initially by letter and by a follow-up telephone call. Participants were white, middle-class residents of a southwestern university town. All were native speakers of American English. Four female and 4 male infants participated in each age group. All mothers participated voluntarily and were not compensated for their participation.

Materials

Seven objects were selected to assist in eliciting comparable labels in adult-directed and infant-directed speech. The toys selected were: A 4-in. (10.16 cm) painted wooden fish; a child's paintbrush; a multicolored woven bracelet; a gray rubber puppet shaped like an elephant; a 6-in. (15.24 cm) plastic brown kangaroo; a plastic green alligator 11-in. (27.94 cm) long; and a blue cloth hippopotamus. The items were chosen due to their presumed ease-of-recognition and their differences in the stress patterns of their names. Accordingly, several word lengths and stress patterns are represented, including a monosyllabic word, *fish*, a disyllabic compound word, *páintbrùsh*, a stress-final trisyllabic word, *kàngaróo*, and a two-, a three-, a four- and a five-syllable word, each with a final unstressed syllable (i.e., *brácelet, élephant, álligàtor*, and *hìppopótamus*, respectively). The monosyllabic word permitted comparisons with previous research using monosyllables, and the stress-final words permitted an analysis of the combined effects of stress and position. The words ending in unstressed syllables were of particular interest because they permitted an assessment of the prediction that unstressed final syllables would be highlighted in infant-directed speech.

The items represented objects that infants in this age range might previously have seen, but the objects were not likely to be a usual play item. For example, a paintbrush was chosen because it is easily recognizable as such, but most 6- and 9-month-olds have not yet played with one. This level of familiarity was chosen so that parents would be motivated to label the object for the child, but so that the object would not be so unusual that parents would label it inconsistently or use abnormally exaggerated intonation.

Procedure

Prior to the experiment, mothers were told only that the study concerned infants' early verbal behavior. Participants and their infants were recorded at their home for approximately 1 hour by two female researchers. One researcher was responsible mainly for interviewing the parent and for structuring the observation. The second researcher operated the audio equipment and played with the infant while the

parent answered the researcher's questions. Participants were audio-recorded using either a Sony TCS-430 cassette recorder or a General Electric AVR-3-5366A cassette recorder. A microphone was placed close to the mother so that it was between the mother and the researcher. When the initial interview was completed, the second researcher typically handed the infant to the mother. Thus, the microphone remained in the same position, relative to the mother, for both the adult-directed and infant-directed portions of the session. The entire session was recorded to allow for later speech analysis.

Following introductory conversation and the signing of the consent form, the mother was asked to interact with her infant per their usual schedule. During a portion of this period, a structured interview was administered. The interview contained questions concerning the infant's daily routine and verbal behavior, and it provided an opportunity for the mother to become comfortable with the researchers and tape recorder.

After this preliminary interaction, mothers were asked to participate in a labeling game. The researcher explained that the game involved two parts: First, the parent would be shown a set of seven objects and asked to identify each object; second, the parent would be asked to introduce each item to the infant. To begin, the researcher showed each of the seven items to the parent one at a time. After the adult-directed labeling, the toys were put into a pile, and the parent was asked to introduce each object, individually, to the infant. Parents were encouraged to allow their infant to play with any of the toys and to take their time. The labeling portion of the session lasted approximately one-half hour.

If the parent used a different label when showing the object to the infant than the label given to the researcher, she was asked to provide that second label to the researcher. This was accomplished by showing the item to the parent and asking, "What name did you give this item?"

Coding

Infant-directed and adult-directed variants of each target word were extracted from the audiotaped speech and digitized at 44.1 kHz using the SoundAccess program on a Macintosh IIci equipped with an Audiomedia sound board. Individual soundfiles were created for each word. These files were then analyzed using Signalyze, a speech analysis package. Duration for the whole word and for each syllable within the word was measured (in milliseconds) from the raw signal, with syllable boundaries being identified both visually and auditorily. To measure pitch peaks, pitch was extracted using a fast fourier transform combination (FFT-Comb) routine in Signalyze. In this routine, successive 512-point FFTs are carried out and the fundamental frequency (F_0) contour is defined by combining frequency information from F_0 and its harmonic frequencies. Pitch peaks were identified visually and were measured from the F_0 contour produced by the FFT-Comb routine. Amplitude was measured by calculating a root mean square (RMS) envelope for each signal. The RMS values were calculated within a window of 50 ms. Measures of peak amplitude for the whole word and for each syllable were taken from the RMS envelope. These measures were obtained in system-based units of linear amplitude, that is, the units used by the computer program to compute an amplitude envelope. Though not standard units, these measures are provided in tables for the purpose of comparisons within this study. For relative measures, decibel differences were calculated as described later.

As a check on the reliability of acoustic analyses, a second coder analyzed a randomly selected 20% of the total word and final syllable samples. At least one of the two coders for a given sample was blind to the type of speech and to the hypothesis of the experiment. The average agreement, obtained by correlating the measures achieved by the two raters, was high, (r = .98), combined across all measures (infant- and adult-directed duration, pitch, and amplitude for whole words and final syllables). The range was r = .89 to r = 1.00; with the exception of pitch in adult-directed speech (r = .89 and r = .90 for whole word and final syllable measures, respectively), rs were between .99 and 1.00. For the analysis of utterance-medial targets, reliabilities were calculated on 17% of the samples, and they were very high, averaging 1.00 across measures of duration, pitch, and amplitude for whole words and final syllables.

RESULTS

In the initial set of analyses, virtually all target words were in final position in an utterance. The reason for analyzing utterance-final words is that, due to the nature of the labeling task, target words were only very rarely produced in other than utterance-final position in adult-directed speech. To permit comparisons across adult- and infant-directed speech, therefore, target words in utterance-final locations were used. Although target words rarely occurred in other than utterance-final position in adult-directed speech and usually occurred in utterance-final position in infant-directed speech, there were some occurrences in infant-directed speech of utterance-medial targets. The frequency of utterance-medial productions varied greatly across mothers, ranging from 0 to 13. The utterance-medial targets were analyzed in a separate set of analyses. Target words generally were not produced in isolation. If a target word was produced multiple times, the first production was normally analyzed, unless the first instance was unintelligible.

Infant-directed and adult-directed samples were always matched in the data submitted to analysis; in three cases, either an infant-directed or an adult-directed sample was not available for a given participant and for a given target word (this occurred once each for *bracelet, fish*, and *alligator*). In those cases, the remaining unpaired sample was excluded from analysis. In addition, participants sometimes used a word that was different from the intended target (e.g., *croc* for *alligator, circle* for *bracelet*). Those samples were also excluded from analy-

sis. In the case of *hippopotamus*, however, most parents used a word that was different from our intended label; as a result, we decided to analyze the most frequently used word, *hippo*, rather than the intended label. For the analyses by syllable type, we used both versions. After exclusions, the numbers of participants contributing data were 13 for *alligator*, 11 for *bracelet*, 16 for *elephant*, 14 for *fish*, 16 for *kangaroo*, 12 for *paintbrush*, and 9 for *hippo* (there were 6 for *hippopotamus*).

Analyses by Word-Level Stress Pattern

An initial set of analyses examined characteristics of final syllables. Because of differences in the stress patterns of the target words, three separate sets of analyses were carried out. A first set of analyses compared the duration, pitch peak, and peak amplitude of the monosyllabic word, *fish*, in infant- and adult-directed speech. A second analysis compared durations, pitch peaks, and peak amplitudes for the two speech types in the words with stress on the final syllable, *kangaroo* and *paintbrush*. A third set of analyses compared these properties in infant- and adult-directed speech for words containing unstressed final syllables (i.e., *alligator, bracelet, elephant*, and *hippo*).

Paired *t* tests were used for these first two sets of comparisons. For the monosyllabic word, separate *t* tests compared the whole-word duration, pitch peak, and peak amplitude across the two speech types. The duration of *fish* in infant-directed speech was significantly longer than in adult-directed speech, $t(13) = 2.96$, $p < .02$. The other two comparisons failed to reach significance. Mean values for each of these measures are presented in the upper sections of Tables 1 through 3.

For the final-stress words, several paired *t* tests were conducted, each for duration, pitch, and amplitude. The duration comparisons consisted of whole-word duration, final-syllable duration, and the proportion of duration accounted for by the final syllable (calculated by taking the final-syllable duration and dividing it by the whole-word duration). Mean scores for these measures are shown in the middle section of Table 1. All three of these comparisons were significant: In infant-directed speech, as compared to adult-directed speech, whole words and final syllables were significantly longer, $t(15) = 3.15$, $p < .005$, and $t(15) = 3.48$, $p < .005$, respectively, and final syllables accounted for a relatively greater amount of the duration, $t(15) = 2.61$, $p < .025$.

The measures analyzed for pitch in words with final stress included the whole-word pitch peak (i.e., the maximum pitch in the word as a whole), nonfinal pitch peak (i.e., the maximum pitch in the word excluding the final syllable), the final-syllable pitch peak, as well as two measures of relative pitch, a pitch difference measure (i.e., the final peak subtracted from the nonfinal peak), and a pitch change measure (the pitch difference divided by the nonfinal pitch peak, to correct for any overall differences in pitch between infant- and adult-directed speech samples). Means for these measures are shown in the middle portion of Table 2. Significant differences between adult-directed and infant-directed speech were observed for whole-word pitch peak, $t(15) = 3.47$, $p < .005$, final-syllable pitch peak, $t(15) = 3.61$, $p < .005$, pitch difference, $t(15) = 3.22$, $p < .01$, and pitch change, $t(15) = 3.03$, $p < .01$. In comparison to adult-directed speech, pitch peaks were higher for whole words and final syllables in infant-directed speech, and both the absolute differ-

TABLE 1
Mean Duration (in ms) of Target Words in Adult-Directed and Infant-Directed Speech

	Whole Word	Final Syllable	Final/Word
Monosyllabic Targets			
Adult-Directed Speech	392	—	—
Infant-Directed Speech	546	—	—
Final-Stressed Targets			
Adult-Directed Speech	667	358	.53
Infant-Directed Speech	877	530	.58
Final-Unstressed Targets			
Adult-Directed Speech	532	212	.40
Infant-Directed Speech	721	346	.47

TABLE 2
Mean Peak Pitch (in Hz) of Target Words in Adult-Directed and Infant-Directed Speech

	Whole Word	Nonfinal	Final	Difference	% Change
Monosyllabic Words					
Adult-Directed Speech	375	—	—	—	—
Infant-Directed Speech	414	—	—	—	—
Final-Stress Words					
Adult-Directed Speech	350	325	340	-15.7	-.06
Infant-Directed Speech	403	316	399	-83.1	-.29
Final-Unstressed Words					
Adult-Directed Speech	363	336	344	-8.1	-.03
Infant-Directed Speech	428	390	355	34.3	.07

TABLE 3
Mean Peak Amplitude[a] of Target Words in Adult-Directed and Infant-Directed Speech

	Whole Word	Nonfinal	Final	dB Difference
Monosyllabic Words				
Adult-Directed Speech	2880	—	—	—
Infant-Directed Speech	4828	—	—	—
Final-Stress Words				
Adult-Directed Speech	3142	3005	2344	1.7
Infant-Directed Speech	4346	3469	4032	-1.1
Final-Unstressed Words				
Adult-Directed Speech	3975	3858	2723	3.1
Infant-Directed Speech	5445	5248	4069	2.7

[a]With the exception of amplitude difference, which is in decibels, all means are in system-based units of linear amplitude obtained from the RMS envelope; these means are provided only for the purpose of comparisons within this study.

ence and the relative difference between final and nonfinal pitch was greater.

For amplitude, differences in whole-word peak amplitude, nonfinal peak amplitude, and final-syllable peak amplitude were analyzed using the system-based units of linear amplitude obtained from the RMS envelope. To obtain a measure of relative amplitude, a decibel difference was computed by taking the log of the ratio of nonfinal peak amplitude to final peak amplitude and multiplying it by 20. Means for whole-word peak amplitude, nonfinal peak amplitude, final peak amplitude (in system units), and amplitude difference (in decibel units) are presented in the center portion of Table 3. The final syllable and decibel-difference comparisons achieved significance, $t(15) = 2.63$, $p < .02$, and $t(15) = 4.11$, $p < .002$, respectively. The whole-word amplitude peak comparison was marginal, $t(15) = 1.93$, $p \approx .073$, and the nonfinal peak amplitude comparison did not achieve significance. Not only were stressed final syllables significantly louder in infant-directed as compared to adult-directed speech, but they were louder relative to the overall amplitude of the word.

The analyses of greatest interest for our hypotheses were the analyses of words containing unstressed final syllables. For these analyses, repeated-measures analyses of variance were carried out separately for duration, pitch, and amplitude. Age and sex were between-subjects factors and speech type (adult-directed vs. infant-directed) was a within-subjects factor. Analyses were conducted on mean scores calculated separately, across words, for duration, pitch, and amplitude and for infant-directed and adult-directed speech. For duration, separate ANOVAs were carried out for whole-word duration, final-syllable duration, and for the proportion of duration accounted for by the final syllable. Means for these duration measures are shown in the bottom portion of Table 1. Significant main effects of speech type (i.e., adult-directed vs. infant-directed) were observed for whole-word duration, final-syllable duration, and for proportion final-syllable duration, $F(1, 12) = 26.13$, $p < .001$, $F(1, 12) =$

26.51, $p < .001$, and $F(1, 12) = 8.47$, $p < .02$, respectively. Duration is longer in infant-directed speech than in adult-directed speech, and this effect is particularly pronounced for final syllables; unstressed final syllables account for a greater proportion of the duration of a word in infant-directed speech.

In addition to the expected effects of speech type, the main effect of sex was significant for final-syllable duration, $F(1, 12) = 7.74$, $p < .05$; a similar trend was observed for whole-word duration, $F(1, 12) = 4.62$, $p \approx .053$. The main effect of sex is due to longer durations in speech of mothers of girls. This effect is observed both in adult-directed and in infant-directed speech. However, the sex × speech type interaction was not significant, suggesting that the effects of speech type were similar across both sexes, that is, despite the slower rate of speech among mothers of girls, the differences between adult- and infant-directed speech were similar across both sexes. No other main effects or interactions were significant.

For analyses of pitch, whole-word pitch peak, nonfinal pitch peak, final-syllable pitch peak, pitch difference, and pitch change were calculated as described previously for final stressed words, and separate ANOVAs were calculated for each measure. Means for these measures are provided in the bottom section of Table 2. Pitch peaks generally were higher in infant-directed than in adult-directed speech both for whole words, $F(1, 12) = 23.28$, $p < .001$, and for nonfinal pitch peak, $F(1, 12) = 25.41$, $p < .001$. The effect of speech type was not significant for final pitch peak, and it was only marginal for pitch difference, $F(1, 12) = 3.56$, $p \approx .084$, and for pitch change, $F(1, 12) = 3.29$, $p \approx .095$.[1] A sex × type interaction was observed for nonfinal pitch peak, $F(1, 12) = 4.83$, $p < .05$; mothers of boys used higher pitch in adult-directed speech than did mothers of girls, whereas mothers of girls produced higher pitched infant-directed speech. Because of this interaction, the adult- versus infant-directed difference in peak pitch for nonfinal syllables was significant only for mothers of girls, $t(7) = 5.09$, $p < .002$; the difference was only marginally significant among mothers of boys, $t(7) = 2.24$, $p \approx .061$. No other main effects or interactions were significant.

The amplitude comparisons included analyses of whole-word peak amplitude, nonfinal peak amplitude, and final-syllable peak amplitude. As described earlier for final-stress words, these analyses were based on system-based units of linear amplitude. Relative amplitude differences were analyzed using decibel units, calculated as described earlier. Means for these measures are presented in the bottom section of Table 3.

Not only was amplitude higher in infant-directed than in adult-directed speech, but it remained relatively higher across the final syllable in infant-directed speech: Whole-word peak amplitude, nonfinal peak amplitude, and final peak amplitude were significantly higher in infant-directed than in adult-directed speech, $F(1, 12) = 6.96$, $p < .025$, $F(1, 12) = 7.18$, $p < .025$, and $F(1, 12) = 5.07$, $p < .05$, respectively. The effect of speech type was nonsignificant for amplitude difference. No other main effects or interactions were significant for any of these four measures.

Consistency Across Participants

The effects of duration were relatively consistent across mothers, with 15 of 16 mothers showing increased duration on final syllables, and 12 of 16 showing greater duration on the final syllable relative to the duration of the word as a whole. These numbers were significant using the binomial test, with $p < .001$ and $p \approx .038$, respectively. Mothers were remarkably consistent in using higher pitch in their infant-directed speech: Whole-word pitch peaks were higher in infant-directed than in adult-directed speech for all 16 mothers, $p < .001$. For final syllables, only 8 of 16 mothers showed higher pitch peaks in infant-directed speech. Interestingly, 13 of 16 mothers showed a greater pitch change in infant-as compared to adult-directed speech, $p \approx .011$; however, there was variability in whether the pitch change reflected an increase or decrease. The heightened amplitude in infant-directed speech was also relatively consistent, with 13 of 16 mothers showing higher peak amplitude both for whole words and for final syllables, $p \approx .011$. No consistency was observed for size of

[1] Because the dependent variable for the pitch change analysis was expressed in proportions, analyses were carried out with the data transformed using an arcsine transformation (Neter, Wasserman, & Kutner, 1985). The pitch change measure was nonsignificant regardless of whether this transformation was performed.

amplitude change between infant- and adult-directed speech.

Word-by-Word Analyses

In addition to the analyses conducted for each word-level stress pattern, individual paired *t* tests were carried out to compare adult-directed with infant-directed speech on a word-by-word basis. These *t* tests were carried out separately for duration, pitch peaks and peak amplitude, and final syllables and whole words. The results of the word-by-word analyses generally were consistent with the results of the previous analyses. For every word, both whole-word and final-syllable duration was greater in infant-directed than in adult-directed speech; this difference was significant ($p < .05$, two-tailed) in both analyses for every word except *hippo*. Whole-word peak pitch was higher for infant-directed than for adult-directed speech in every case, but reached significance only for *elephant* and *kangaroo*. Final-syllable pitch peak was less consistently higher in infant- than adult-directed speech, though it was significantly higher for *hippo* and *kangaroo*. Peak amplitude was higher in each whole word except *paintbrush*, and final-syllable peak amplitude was higher in every word; however, the differences were significant only for *hippo* and *kangaroo* in both analyses. Thus, although the effects for duration are quite robust, the peak pitch and amplitude differences do not reach significance for most of the individual words. Nonetheless, the general pattern of greater duration, higher pitch, and heightened peak amplitude in infant-directed speech is quite consistent across words, with the exception that pitch is not typically higher for final syllables.

Analyses by Syllable Type

To assess more fully the relative contributions of stress and final position to the form of mothers' productions, a separate set of analyses was carried out in which syllables of words in infant- and adult-directed speech were analyzed as a function of stress and position. Specifically, separate means were calculated for syllables categorized as nonfinal stressed, nonfinal unstressed, final stressed, and final unstressed. Primary and secondary stresses were not distinguished in this analysis; both were counted as "stressed" syllables. Because this analysis was based on syllables and not words, the instances of *hippopotamus* were included; the inclusion of *hippopotamus* made little difference either in specific means or in the overall pattern of significant and nonsignificant results. The means for duration, pitch, and amplitude are shown in Tables 4, 5, and 6, respectively.

The means were entered into a multivariate analysis of variance, in which duration, pitch peak, and peak amplitude were the dependent measures and type of speech (infant- vs. adult-directed), stress (stressed vs. unstressed), and position (final vs. nonfinal) were within-subjects factors. The multivariate analysis of speech type was significant, $F(3, 13) = 7.91$, $p < .005$. Each of the univariate F tests was significant, with $F(1, 15) = 21.12$, $p < .001$, $F(1, 15) = 11.91$, $p < .005$, and $F(1, 15) = 10.49$, $p < .01$, for duration, pitch peak, and peak amplitude, respectively. Infant-directed syllables were longer in duration, higher in pitch, and louder than were adult-directed syllables. The multivariate analysis of syllable position was also significant, $F(3, 13) = 27.24$, $p < .001$. Univariate tests for duration and pitch were significant, $F(1, 15) = 71.30$, $p < .001$, and $F(1, 15) = 31.76$, $p < .001$, respectively. The univariate test for amplitude was not significant. Final syllables were longer than nonfinal syllables, and they were also higher pitched. For stress, the multivariate analysis reached significance, $F(3, 13) = 20.22$, $p < .001$, and the univariate analyses were significant for duration and pitch, $F(1, 15) = 54.33$, $p < .001$, and $F(1, 15) = 23.12$, $p < .001$, respectively. The univariate test for amplitude was again nonsignificant. Stressed syllables were longer and higher pitched than unstressed syllables.

In analyses of interactions, the type × syllable position interaction was marginally significant at

TABLE 4
Mean Duration (in ms) in Adult-Directed and Infant-Directed Speech for Different Stress Levels and Positions

	Nonfinal	Final
Adult-Directed Speech		
Stressed	206	357
Unstressed	134	229
Infant-Directed Speech		
Stressed	348	530
Unstressed	138	355

TABLE 5
Mean Pitch Peaks (in Hz) in Adult-Directed and Infant-Directed Speech for Different Stress Levels and Positions

	Nonfinal	Final
Adult-Directed Speech		
Stressed	315	340
Unstressed	309	345
Infant-Directed Speech		
Stressed	343	399
Unstressed	315	360

TABLE 6
Mean Peak Amplitude[a] in Adult-Directed and Infant-Directed Speech for Different Stress Levels and Positions

	Nonfinal	Final
Adult-Directed Speech		
Stressed	3163	2344
Unstressed	2301	2637
Infant-Directed Speech		
Stressed	4196	4032
Unstressed	3983	3962

[a]All means are in system-based units of linear amplitude and are provided only for the purpose of comparisons within this study.

the multivariate level, $F(3, 13) = 3.07$, $p \approx .066$. Among univariate analyses, that for duration was significant, $F(1, 15) = 7.14$, $p < .02$, and that for pitch was marginal, $F(1, 15) = 3.14$, $p \approx .097$. Final syllables in infant-directed speech were longer than in adult-directed speech, and they were marginally higher in pitch. The multivariate type × stress analysis was significant, $F(3, 13) = 10.76$, $p < .002$. Univariate analyses were significant for duration, $F(1, 15) = 15.85$, $p < .002$, and pitch, $F(1, 15) = 6.28$, $p < .025$, but not for amplitude. Stressed syllables were longer and higher pitched in infant-directed speech than in adult-directed speech. The position × stress interaction was not significant at the multivariate level, and none of the univariate tests was significant. Thus, where effects of stress and position were observed, those effects were additive, that is, effects did not differ for different levels of stress or position. The type × position × stress interaction was marginal, $F(3, 13) = 3.34$, $p \approx .053$. Among the univariate tests, only that for duration was significant, $F(1, 15) = 6.11$, $p < .05$. This interaction reflects the pattern, apparent in Table 4, wherein syllables that are both unstressed and nonfinal do not differ between infant- and adult-directed speech, whereas those that are either stressed or final do, with the differences being greatest for syllables that are both stressed and final.

To test specifically whether unstressed word-final syllables differed in duration from unstressed nonfinal syllables, paired t tests were conducted employing Bonferroni's correction (as recommended by Myers, 1979). Because there were 12 comparisons of theoretical interest (4 each within infant- and adult-directed speech and 4 between-register comparisons), the Bonferroni correction resulted in a per-test α level of .004 for an experiment-wise α of .05. Employing these criteria, each of the pairwise comparisons within the registers of infant- and adult-directed speech were significant (i.e., stressed nonfinal differed from stressed final and unstressed nonfinal; unstressed final differed from unstressed nonfinal and stressed final). All of the between-register pairwise comparisons were significant except that between adult- and infant-directed unstressed nonfinal syllables, $t(15) = 0.42$, *ns*.

Analysis of Utterance-Medial Word-Final Syllables

These analyses were conducted on utterance-final targets because targets most frequently occurred in final position, especially in adult-directed speech. The corpus contained some instances of target words in utterance-medial positions in infant-directed speech, but there were very few examples in adult-directed speech. The absence of a sufficient number of utterance-medial targets in adult-directed speech prevented comparisons between infant-directed and adult-directed speech like those conducted for utterance-final targets. However, it was possible to compare characteristics of final and nonfinal syllables within infant-directed speech.

Only 10 of the original 16 mothers contributed data for the utterance-internal analyses (5 were mothers of 9-month-olds, and 5 were mothers of 6-month-olds). The number of targets identified in utterance-medial position was 41. However, if a target word was produced multiple times in similar syntactic contexts (e.g., "What's the elephant do?" "What's that elephant doing?"), only the first instance was analyzed. In addition, productions of the mono-

syllabic target were not included in this analysis. After these exclusions, 35 words were analyzed. Of these, 28 words had unstressed final syllables, and 7 had stressed final syllables. Because target words rarely occurred in phrase-internal locations, most of the analyzed words occurred in phrase-final position; only 4 were not phrase-final. Although it would have been interesting to analyze more words in non-phrase-final locations, nouns in English generally occur in phrase-final position, and, thus, the distribution of phrase-final and non-phrase-final productions in this sample is probably fairly representative. Four of the targets occurred in clause-final position.

For the purpose of analysis, individual syllables were categorized as stressed or unstressed and as word-final or non-word-final. Separate means were calculated, for each participant, for the duration, peak pitch, and peak amplitude of syllables categorized as nonfinal stressed, nonfinal unstressed, final stressed, and final unstressed. The results were analyzed using paired *t* tests. Because analyses required multiple comparisons on the same data set, the Bonferroni test was employed. Only 5 participants produced any words containing stressed final syllables in utterance-medial position. Accordingly, no comparisons were conducted with stressed final syllables. For the two comparisons remaining, the Bonferroni correction results in a per-test α level of .025 for a familywise error rate of .05. In analyses of duration, unstressed final syllables were longer than unstressed nonfinal syllables, $t(9) = 3.20$, $p < .02$, and stressed nonfinal syllables were longer than unstressed nonfinal syllables, $t(9) = 4.10$, $p < .005$. The mean duration values for utterance-medial syllables were 199 ms for nonfinal stressed, 245 ms for final stressed, 125 ms for nonfinal unstressed, and 196 ms for final unstressed syllables. For pitch and amplitude, neither of the two comparisons approached significance.

DISCUSSION

The results of the whole-word analyses support previous suggestions that infant-directed speech is slower and higher pitched than is adult-directed speech (e.g., Fernald & Simon, 1984; Garnica, 1977). In analyses of monosyllabic, final-stressed, and final-unstressed targets, words were consistently longer in infant-directed speech as compared to adult-directed speech. Pitch peaks were significantly higher for both types of multisyllabic targets. Additionally, amplitude peaks were greater in infant-directed speech for targets with unstressed final syllables and were marginally greater for targets with stressed final syllables. Thus, the results extend beyond the prior findings in suggesting that some types of infant-directed speech may be louder than the comparable adult-directed speech.

The analyses of word-final syllables revealed, as predicted, that word-final syllables in the infant-directed samples were of greater duration than final syllables in the adult-directed samples. Duration of final syllables was enhanced even relative to the overall increase in duration observed in infant-directed speech: Final syllables accounted for a greater proportion of the whole-word duration in infant-directed than in adult-directed speech. This result was observed both for stressed and for unstressed final syllables. Caregivers may indeed elongate words as they speak to their infants, and they may particularly exaggerate the length of final syllables. This added length may in turn contribute to the perceptual salience of final syllables for the young language learner.

The heightened pitch in infant-directed speech extends to final syllables in our sample, but only for words ending in stressed syllables: Unstressed final syllables are not significantly higher pitched in infant-directed than in adult-directed speech. This finding is not necessarily surprising: Given that infant-directed speech typically exhibits a greater pitch range than adult-directed speech (e.g., Fernald & Mazzie, 1991; Fernald & Simon, 1984; Garnica, 1977), final unstressed syllables may often fall in the lower part of the pitch range. Though our data do not provide direct evidence of a greater pitch range in infant-directed speech, we analyzed only the maximum pitch level within a syllable; an assessment of pitch range would require measuring pitch minima as well as maxima within each syllable. Moreover, the finding that mothers tended to show greater pitch changes in infant- as opposed to adult-directed speech, even though the direction of change varied, is consistent with the notion that pitch fluctuates more in infant-directed speech.

Peak amplitude was higher in infant-directed speech than in adult-directed speech for final syllables both for targets containing stressed final syllables and for those containing unstressed final syllables. Higher amplitude has not frequently been described as one of the properties of infant-directed speech. This observation is, however, consistent with the notion that infant-directed speech is serving an attention-getting or highlighting function. The interpretation of amplitude differences is somewhat tricky, however, because these measures are highly sensitive to such artifacts of the recording context as placement of the microphone and distance of the speaker from the microphone.[2] Because of the potential difficulty in interpreting amplitude data, the relative measures of amplitude are particularly useful. The amplitude difference (calculated in decibels) does differ between infant-directed and adult-directed speech for targets with stressed final syllables, but not for targets with unstressed final syllables. Thus, mothers appear to be exaggerating the loudness of stressed final syllables but not of unstressed final syllables.

Based on the results of the analyses of final syllables, we might ask about the nature of the highlighting in infant-directed speech. One possibility is that mothers are putting special emphasis on final syllables in infant-directed speech, with the result that final unstressed syllables are similar to stressed syllables in this register; in adult-directed speech, final unstressed syllables may be more similar to nonfinal unstressed syllables. Alternatively, it could be that all syllables are emphasized and are therefore more similar to one another in infant-directed speech. In that event, the utility of stress cues for segmentation would be minimal. The finding that final syllables represent a greater proportion of the length of a word in infant-directed speech as compared to adult-directed speech argues against the possibility that all syllables are emphasized to a similar degree in infant-directed speech: If it were the case that even unstressed nonfinal syllables exhibited increased duration, then it is unlikely that final syllables could be relatively longer in the infant-directed register. Nonetheless, the possibilities raised here suggested a need to examine the characteristics of unstressed nonfinal syllables.

When the results of the analyses by syllable type are examined, it becomes apparent that all types of syllables are not more similar to one another in infant-directed speech, nor are final syllables uniquely highlighted. Instead, particularly with respect to duration but also to some extent with pitch, highlighting that is present in adult-directed speech is greatly exaggerated in infant-directed speech. Whereas the duration of unstressed, nonfinal syllables does not differ between infant- and adult-directed speech, the duration of stressed and final syllables is much greater in the infant-directed register, and that of syllables that are both stressed and final is particularly exaggerated. For pitch, there is a tendency for syllables that are emphasized in adult-directed speech to receive enhanced emphasis in infant-directed speech. This tendency is most pronounced for stressed syllables, however, and does not appear to characterize unstressed final syllables: There was some tendency to increase pitch on final syllables both in adult-directed and infant-directed speech, but this increase was not greater in infant-directed as compared to adult-directed speech.[3] For amplitude, on the other hand, differences between syllables appear compressed in infant-directed speech. Although the amplitude change measure differed significantly between registers for final stressed syllables in the first set of analyses, which

[2] It is unlikely that these factors biased the amplitude results: The microphone was normally in a stationary location with respect to the mother throughout the experiment. If any bias were present, it would tend to favor increased amplitude in adult-directed, not infant-directed, speech: For adult-directed samples, the mother was speaking directly toward the microphone and toward a researcher who was a short distance away. For infant-directed samples, the mother would have been speaking to a child on her lap, so she would be closer to the listener and turned away from the microphone.

[3] There are two possible reasons for the increased pitch on final syllables in the adult-directed samples: First, mothers sometimes used rising intonation on their utterances during the labeling task, using prosody akin to list intonation, even though separate utterances were used to label the objects. Second, some of the mothers, even though talking to an adult, used a register approaching a motherese register (presumably because their infant was present). Both tendencies should either be neutral or should work against our hypotheses with respect to duration: The rising intonation contour, if it has any effect, would tend to increase duration of final syllables in adult-directed speech, thus making it less likely that length differences between infant- and adult-directed speech would be observed. The use of a motherese register for some adult-directed speech would also tend to decrease differences between the two samples.

suggested that final stressed syllables are especially loud in infant-directed speech, the results of the analyses by syllable type suggest a different picture: The difference between registers is due to the relatively low amplitude of final stressed syllables in adult-directed speech more than to a relatively high amplitude on these syllables in infant-directed speech. The effect of speech type coupled with the failure of any other effects or interactions to reach significance for amplitude support the pattern exhibited in Table 6: Syllables in infant-directed speech are consistently louder than in adult-directed speech, but there is little variation as a function of stress or position in infant-directed speech.

These analyses used utterance-final words. The predominance of utterance-final targets in our corpus was a natural consequence of the way in which the target words were elicited, specifically, in a labeling context. Words tend to be placed in utterance-final position in labeling contexts, both in infant- and adult-directed speech (Fernald & Mazzie, 1991). In defense of the findings with utterance-final targets, the increased salience of word-final syllables would be valuable for the initial identification of words even if limited to utterance-final contexts: Focused words are especially likely to be placed in utterance-final position in labeling contexts in infant-directed speech (Fernald & Mazzie, 1991; Woodward & Aslin, 1990). Labeling contexts will tend to be frequent, particularly as children begin to learn language. Moreover, it is in the labeling context that children may be most motivated to identify words from the speech stream. Nonetheless, the difficult segmentation problem is going to be when words are utterance-internal, not when they are utterance-final. Any available evidence of word-final lengthening in utterance-medial positions is therefore of interest.

For standard adult speech, there is substantial evidence that syllables are lengthened in phrase- and clause-final positions within sentences (e.g., Klatt, 1976; Nakatani, O'Connor, & Aston, 1981). Bernstein Ratner (1986) and Morgan (1986) showed that final lengthening is exaggerated in infant-directed speech in phrase- or clause-final position in at least some sentence-internal locations. However, both studies examined lengthening in monosyllabic words and, unlike our research, did not examine unstressed syllables.[4] Moreover, given that these researchers were seeking to identify the relative amount of prephrase or pre-clause-boundary lengthening, they were comparing final syllables across two different contexts and were not comparing final with nonfinal syllables. Thus, our results with utterance-final targets extend this prior research by documenting exaggerated clause-final lengthening in unstressed syllables in infant-directed speech. However, the results of the analyses reported so far do not tell us whether word-final syllables would also be highlighted when they are not in utterance-final position or when they occur in nonlabeling speech.

The analyses of utterance-medial word-final syllables were conducted to ascertain whether exaggerated word-final lengthening generalizes to non-utterance-final contexts. The results of these analyses suggest that there is generality: Unstressed word-final syllables were significantly longer than unstressed non-word-final syllables, even in utterance-medial position. These results are thus consistent wth the proposal that unstressed word-final syllables are more salient than unstressed non-word-final syllables due to increased length, regardless of whether the word occurs at the end of an utterance or internal to an utterance. Nonfinal stressed syllables are also longer than nonfinal unstressed syllables, a result consistent with earlier analyses suggesting that stressed syllables are of greater duration than unstressed syllables. The failure to find any effects of pitch or amplitude could suggest that duration plays a greater role in highlighting salient syllables in our sample than do these other acoustic properties; however, given the small number of participants and the limited number of utterances contributing to this analysis, these null results must be treated with caution.

Though present, word-final lengthening is much more limited in utterance-internal locations than in utterance-final position. The utterance-medial lengthening could be more limited because syllables in utterance-final position benefit both from word- and sentence-final

[4] Morgan (1986) analyzed only content words, so his targets should all have been stressed; Bernstein Ratner (1986) did not mention including or excluding any particular types of words but, given that her targets fell in clause-final position, it is likely that most were content words.

lengthening, and both types of lengthening are exaggerated in infant-directed speech. Alternatively, words in utterance-medial position were, in most cases, no longer "new." Target words usually were produced in utterance-medial position only after those words were introduced at least once in final position. Evidence from standard adult speech suggests that "new" words are lengthened relative to previously presented words (Fowler & Housam, 1987).

Some cautions are in order here. Although our data provide clear evidence that final syllables in this sample of infant-directed speech are longer than their adult-directed counterparts, this sample was restricted to a context in which parents were introducing novel toys to their children. The objects were chosen to be sufficiently common that a child could have been exposed to the objects, but some objects may have been labeled for the child only infrequently, and the particular exemplars were, in any event, new to the child. Focused words are particularly likely to be highlighted through placement on pitch peaks in infant-directed speech (Fernald & Mazzie, 1991), and they are lengthened in adult-directed speech (Fowler & Housam, 1987). The data from the utterance-internal targets alleviate this problem somewhat because the targets contributing to that analysis usually were not "new." Nonetheless, the context was restricted. In addition to the restricted context, our data are limited to English and, specifically, to middle-class English-speaking mothers. Moreover, the target words were all nouns. We do not know whether similar results would extend to verbs or adjectives. However, given that object terms tend to predominate in the early vocabularies of English-learning children (Gentner, 1982; Nelson, 1973), the segmentation of nouns will be particularly important for early word learning.

In combination, the results of the analyses of utterance-final and utterance-internal targets provide additional evidence of ways in which prosody may facilitate word learning. The heightened pitch and amplitude, and the increased duration, may help to elicit the infant's interest; the final lengthening may serve to hold the infant's attention. These characteristics of infant-directed speech may also serve to increase the salience of target words by focusing the child's attention on the highlighted syllables. This may hold true even for final syllables that are normally unstressed in English, given our finding that even unstressed final syllables are relatively longer in infant- than adult-directed speech. Combined with the exaggerated salience of stressed syllables, the addition of stress cues such as increased duration to unstressed final syllables could provide a valuable tool for word-level segmentation, enabling the child to extract both a stressed and an unstressed syllable of a word. Moreover, the increased duration on final syllables would explain suggestions in the literature that word-final syllables, as well as stressed syllables, are perceptually salient to young children and are preferentially extracted, stored, and included in early productions (e.g., Echols & Newport, 1992; Peters, 1983). If word-final syllables frequently carry certain characteristic of stress, then it may be that the retention both of stressed and final syllables can be accounted for by a single tendency to attend to "stressed" syllables.

Although the observation that final syllables carry some characteristics of stress in infant-directed speech could help to provide an acoustic basis for prior suggestions that final syllables, like stressed syllables, are perceptually salient to the young language learner, the utility of this salience for word-level segmentation may appear limited for several reasons. First, a great many of the words that young children need to identify will be monosyllabic words. Lengthening of final unstressed syllables will not help in the segmentation of these words. On the other hand, monosyllabic content words are stressed. If stressed syllables are particularly salient and are readily extractable, then the segmentation of monosyllabic content words should not be problematic. Moreover, the phrase- and clause-final lengthening observed for monosyllabic words in prior research (Bernstein Ratner, 1986; Morgan, 1986) and replicated here for monosyllables in utterance-final position, should further assist in the identification of these words. Though multisyllabic words are less frequent in English, they constitute the category of (content) words for which the segmentation problem is more difficult.

A second possible limitation to the utility of word-final lengthening in infant-directed speech is that if this lengthening resulted in the perception of these syllables as stressed, then it could potentially increase, not decrease, the difficulty of segmenting words from speech:

The lengthening would be useful only if the target word were followed by an unstressed syllable, but not when it was followed by a stressed syllable. Although content words will often be followed by a function word, and therefore presumably an unstressed syllable, there are many cases in which a word-final syllable will be followed by a stressed syllable as, for example, in adjective–noun or noun–verb combinations. Because our data include only nouns, we do not know whether the final lengthening would occur for adjectives. However, noun–verb combinations will no doubt be common in infant-directed speech. In such contexts, the treating of unstressed final syllables as stressed could result in the perception of a sequence of stressed syllables, and no segmentation advantage.

The possibility that the highlighting will result in sequences of "stressed" syllables is based on the assumption that lengthened final syllables will be perceived as stressed. However, in suggesting that final syllables in infant-directed speech carry some characteristics of stress, we do not intend to imply that they are indistinguishable from syllables that are underlyingly stressed. The special highlighting that final syllables in infant-directed speech receive is conveyed primarily by duration. In fact, the results of the multivariate analyses for location × type and location × stress × type, wherein the multivariate analyses achieved only marginal significance whereas the univariate analysis for duration was significant, is consistent with this distinction. Although arguing from null results requires caution (especially when the effects were so close to significance), the difference between the univariate results and the multivariate results may suggest that the special highlighting on final syllables is uniquely provided by duration and is not the result of multiple stress cues.[5]

[5] The importance placed on duration in our account may appear counter to prior research suggesting that pitch, not duration, is the aspect of infant-directed speech that is most attractive to infants (Fernald & Kuhl, 1987). Even if one aspect of infant-directed speech is particularly attractive, infants could still notice other characteristics. There may also be age differences: The subjects for Fernald and Kuhl's (1987) study were 4 months old; by 6 to 9 months of age, infants may be increasingly attentive to various properties of speech. Infants as young as 6 months can use rhythmic patterns signaled solely by duration to identify units in speech (Morgan & Saffran, 1995).

The issue of a distinction between cues indicating stress and those highlighting final syllables brings us to the second potential function of stress in infant-directed speech, that of signaling the metrical structure of the native language. If, as has been suggested (e.g., Echols et al., in press), children initially identify the trochaic tendency of English by extracting trochaic words, then the highlighting of unstressed final syllables through increased duration should be useful: Given that the trochaic disyllable is a highly common multisyllabic word type in English and given the trochaic nature of many "babytalk" variants of otherwise monosyllabic words (e.g., *mommy, daddy, birdie*), trochaic disyllables are likely to predominate among multisyllabic words in infant-directed speech. Because both the initial stressed syllable and the lengthened final syllable will be salient, children should find it easy to extract disyllabic words with trochaic stress. Once a number of these disyllables are identified, a child could begin to recognize the characteristic rhythm pattern of these sequences.

On the other hand, our observation that unstressed final syllables carry some characteristics of stress could complicate the task of identifying a trochaic accent pattern, as it is then less clear which cues are determining the trochaic rhythm of the disyllable. Although our results do not provide clear evidence as to the exact cues indicating the trochaic accent pattern, it is likely that a combination of cues typically associated with stress in English (e.g., pitch, duration, vowel quality) could indicate stress. Thus, the cues contributing to extractability and the cues contributing to the identification of a metrical representation for English would differ somewhat: Whereas increased duration on the final syllable would be sufficient to facilitate extraction, multiple stress cues would convey the trochaic nature of the sequence. A complete account of the cues to metrical structure will require a more detailed assessment of stress cues across different contexts. We did not assess vowel quality, and, because we measured only pitch maxima, our measures may not fully reflect pitch changes. In a more extensive analysis, we would expect that stress should be indicated by a correlated set of cues across different contexts, whereas final syllables should be consistently longer in duration.

As was noted earlier, the identification of the trochaic stress pattern is important not only for

the construction of a metrical representation of the native language, but it may itself assist in word-level segmentation. Thus, our account may contribute to an explanation for how rhythm could assist in the beginnings of word-level segmentation. If, as a result of the highlighting described here, prelinguistic infants succeed in extracting a large number of trochaic disyllables, then infants could, potentially, identify the stress pattern common to these disyllabic extractions. That would enable infants to identify the trochaic stress pattern as characteristic of English even before learning many (or any) words of English. After the trochaic stress pattern is identified, an infant could use it to extract additional words from the speech stream. This account could thus explain evidence that prelinguistic infants can use the trochaic stress pattern to extract words from continuous speech (Echols et al., in press; Newsome & Jusczyk, 1995).

This account predicts that infants should find both stressed and final syllables to be salient at an age younger than that at which they expect trochaic disyllables to cohere. Because infants appear capable of using trochaic rhythm for segmentation by 9 months, evidence for the salience of final syllables might be found in even younger infants. The failure to find age differences in the tendency of mothers to increase the duration of final syllables suggests that cues are present in infant-directed speech at least for infants as young as 6 months. The next step will be to determine whether 6-month-olds take advantage of this heightened salience on final syllables, even though infants of this age do not demonstrate a sensitivity to the trochaic stress pattern.

In summary, our results suggest the following: They support prior claims that infant-directed speech has higher pitch peaks than adult-directed speech and that the duration of individual words is greater. They also suggest that words in infant-directed speech are louder. They extend beyond prior research in suggesting that unstressed word-final syllables as well as stressed syllables are highlighted in infant-directed speech. These results may explain evidence that word-final syllables are perceptually salient to young children. They also may contribute to an account for the development of sensitivity to the trochaic stress pattern in English. In so doing, they contribute to an understanding of a fundamental problem of early language development, that of how infants are able to segment word-level units from the stream of speech.

REFERENCES

Allen, G.D., & Hawkins, S. (1980). Phonological rhythm: Definition and development. In G. Yeni-Komshian, J.F. Kavanagh, & C.A. Ferguson (Eds.), *Child phonology: Vol. 1: Production*. New York: Academic.

Bernstein Ratner, N. (1986). Durational cues which mark clause boundaries in mother-child speech. *Journal of Phonetics, 14*, 303–309.

Blasdell, R., & Jensen, P. (1970). Stress and word position as determinants of imitation in first language learners. *Journal of Speech and Hearing Research, 13*, 193–202.

Brown, G. (1983). Prosodic structure and the given/new distinction. In A. Cutler & D.R. Ladd (Eds.), *Prosody: Models and measurements*. Berlin: Springer-Verlag.

Cole, R.A., & Jakamik, J. (1980). A model of speech perception. In R.A. Cole (Ed.), *Perception and production of fluent speech*. Hillsdale, NJ: Erlbaum.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language, 31*, 218–236.

Cutler, A., & Foss, D.J. (1977). On the role of sentence stress in sentence processing. *Language and Speech, 20*, 1–10.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 113–121.

Demuth, K. (1996). The prosodic structure of early words. In J.L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*. Hillsdale, NJ: Erlbaum.

Echols, C.H. (1996). A role for stress in early speech segmentation. In J.L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*. Hillsdale, NJ: Erlbaum.

Echols, C.H., Crowhurst, M.J., & Childers, J.B. (in press). The perception of rhythmic units in speech by infants and adults. *Journal of Memory and Language*.

Echols, C.H., & Newport, E.L. (1992). The role of stress and position in determining first words. *Language Acquisition, 2*, 189–220.

Fernald, A. (1991). Prosody in speech to children: Prelinguistic and linguistic functions. In R. Vasta (Ed.), *Annals of child development* (Vol. 8) London: Kingsley.

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development, 10*, 279–293.

Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology, 27*, 209–221.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mother's speech to newborns. *Developmental Psychology, 27*, 104–113.

Fernald, A., Taeschner, T., Dunn, J., Papoušek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language, 16*, 477–501.

Fikkert, P. (1994). *On the acquisition of prosodic structure*. Dordrecht: ICG Printing.

Fowler, C.A., & Housam, J. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language, 26*, 489–504.

Garnica, O. (1977). Some prosodic and paralinguistic features of speech to young children. In C.E. Snow & C.A. Ferguson (Eds.), *Talking to children: Language input and acquisition*. Cambridge, England: Cambridge University Press.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj (Ed.), *Language development: Language, cognition and culture*. Hillsdale, NJ: Erlbaum.

Gerken, L. (1990). Performance constraints in early language: The case of subject-less sentences. *Papers and Reports on Child Language Development, 29*, 54–61.

Gerken, L.A. (1994a). A metrical template account of children's weak syllable omissions. *Journal of Child Language, 21*, 565–584.

Gerken, L.A. (1994b). Young children's representation of prosodic structure: Evidence from English-speakers' weak syllable omissions. *Journal of Memory and Language, 33*, 19–38.

Gleitman, L.R., Gleitman, H., Landau, B., & Wanner, E. (1988). Where learning begins: Initial representations for language learning. In F.J. Newmeyer (Ed.), *Linguistics: The Cambridge Survey: Vol. 3. Language: Psychological and biological aspects*. Cambridge, England: Cambridge University Press.

Gleitman, L.R., & Wanner, E. (1982). Language acquisition: The state of the art. In L.R. Gleitman & E. Wanner (Eds.), *Language acquisition: The state of the art*. Cambridge: Cambridge University Press.

Grieser, D.L., & Kuhl, P.K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology, 24*, 14–20.

Grosjean, F., & Gee, J.P. (1987). Prosodic structure and spoken word recognition. *Cognition, 25*, 135–155.

Hayes, B. (1985). Iambic and trochaic rhythm in stress rules. *Proceedings of the Berkeley Linguistics Society, 13*, 429–446.

Hayes, J.R., & Clark, H.H. (1970). Experiments on the segmentation of an artificial speech analog. In J.R. Hayes (Ed.), *Cognition and the development of language*. New York: Wiley.

Hirsh-Pasek, K., Kemler Nelson, D.G., Jusczyk, P.W., Cassidy, K.W., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition, 26*, 269–286.

Hura, S.L., & Echols, C.H. (1996). The role of stress and articulatory difficulty in children's early productions. *Developmental Psychology, 32*, 165–176.

Ingram, D. (1978). The role of the syllable in phonological development. In A. Bell & J.B. Hooper (Eds.), *Syllables and segments*. Amsterdam: North-Holland.

Jusczyk, P.W., Cutler, A., & Redanz, N.J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development, 64*, 675–687.

Jusczyk, P.W., Hirsh-Pasek, K., Kemler Nelson, D.G., Kennedy, L.J., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology, 24*, 252–293.

Jusczyk, P.W., & Thompson, E. (1978). Perception of a phonetic contrast in multisyllabic utterances by two month olds. *Perception and Psychophysics, 2*, 105–109.

Kemler Nelson, D.G., Hirsh-Pasek, K., Jusczyk, P.W., & Wright Cassidy, K. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language, 16*, 55–68.

Klatt, D.H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America, 59*, 1208–1221.

Klein, H.B. (1981). Early perceptual strategies for the replication of consonants from polysyllabic lexical models. *Journal of Speech and Hearing Research, 24*, 535–551.

Klein, H.B. (1984). Learning to stress. *Journal of Child Language, 11*, 375–390.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.

Mehler, J., Jusczyk, P.W., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition, 29*, 143–178.

Morgan, J.L. (1986). *From simple input to complex grammar*. Cambridge, MA: MIT Press.

Morgan, J.L. (1996). A rhythmic bias in preverbal speech segmentation. *Journal of Memory and Language, 35*, 666–688.

Morgan, J.L. & Saffran, J.R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development, 66*, 911–936.

Myers, J.L. (1979). *Fundamentals of experimental design* (3rd ed.). Boston: Allyn & Bacon.

Nakatani, L.H., O'Connor, K.D., & Aston, C.H. (1981). Prosodic aspects of American English speech rhythm. *Phonetica, 38*, 84–105.

Nakatani, L.H., & Schaffer, J.A. (1978). Hearing "words" without words: Prosodic cues for word perception. *Journal of the Acoustical Society of America, 63*, 234–245.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development, 38*, (1–2, Serial No. 149).

Neter, J., Wasserman, W., & Kutner, M. (1985). *Applied linear statistical models: Regression, analysis of variance, and experimental designs*. Homewood, IL: Richard D. Irwin, Inc.

Newsome, M., & Jusczyk, P.W. (1995). Do infants use stress as a cue in segmenting fluent speech? In D. MacLaughlin & S. McEwen (Eds.), *Proceedings of the 19th Boston University Conference on Language Development*. Boston, MA: Cascadilla Press.

Peters, A.M. (1983). *The units of language acquisition*. Cambridge, England: Cambridge University Press.

Pollock, K.E., Brammer, D.M., & Hageman, C.F. (1989). An acoustic analysis of young children's productions of word stress. *Papers and Reports on Child Language Development, 28*, 140–147.

Slobin, D.I. (1973). Cognitive prerequisites for the development of grammar. In C.A. Ferguson & D.I. Slobin (Eds.), *Studies of child language development*. New York: Holt, Rinehart, & Winston.

Slobin, D.I. (1985). Crosslinguistic evidence for the language-making capacity. In D. I. Slobin (Ed.), *The cross-linguistic study of language acquisition: Vol 2. Theoretical issues*. Hillsdale, NJ: Erlbaum.

Spring, D.R., & Dale, P.S. (1977). Discrimination of linguistic stress in early infancy. *Journal of Speech and Hearing Research, 20*, 224–232.

Wijnen, F., Krikhaar, E., & den Os, E. (1994). The (non)realization of unstressed elements in children's utterances: Evidence for a rhythmic constraint. *Journal of Child Language, 21*, 59–83.

Woodward, J.Z., & Aslin, R.N. (April, 1990). *Segmentation cues in maternal speech to infants*. Poster presented at the International Conference on Infancy Studies, Montréal, Canada.

13 January 1995; Revised 05 April 1996 ■